

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:36:32 / Search time 46 Seconds
(without alignments)
3114.027 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMWGSFKSLSETCLPN.....ALVLPSTVLDLQLCRYPD 454

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2290	97.4	490	4	Q81UE7
2	2284	97.2	490	4	Q8NPT2
3	2239.5	95.3	489	11	Q8BMB6
4	1318	56.1	488	11	Q8C5F0
5	1316	56.0	488	11	Q8P41
6	1316	56.0	488	11	Q8CIS9
7	1316	56.0	488	11	Q8CIS9
8	1306	55.6	526	11	Q80ZFP3
9	1272	54.1	514	11	Q924Z1
10	1272	54.1	488	4	Q86SF6
11	1269	54.0	498	4	Q72389
12	1252.5	53.0	487	4	Q8NMB5
13	1245.5	53.0	487	4	Q8NMB5
14	1074.5	45.7	456	4	Q8TF03
15	1064.5	45.3	470	11	Q923B6
16	1057.5	45.0	470	11	Q91W31

17	1041.5	44.3	474	11	Q91ZB8	Q91ZB8 mus musculus
18	956.5	40.7	468	11	Q7TPB8	Q7TPB8 ratu
19	728	31.0	338	6	Q9GL50	Q9GL50 sus scrofa
20	715	30.4	339	11	Q924Z2	Q924Z2 mus musculus
21	707	30.1	339	11	Q9CNR7	Q9CNR7 mus musculus
22	704	29.9	339	11	Q924U9	Q924U9 mus musculus
23	628	26.7	283	4	Q9H5R1	Q9H5R1 homo sapien
24	541	23.0	283	4	Q8WMB0	Q8WMB0 homo sapien
25	324	13.8	143	4	Q9H7Y1	Q9H7Y1 homo sapien
26	171.5	7.3	208	16	Q8NLU6	Q8NLU6 corynebacte
27	170	7.2	208	16	Q8XQ52	Q8XQ52 raltosonia s
28	164.5	7.0	234	16	Q82AX0	Q82AX0 streptomyce
29	160	6.8	198	16	Q930X7	Q930X7 rhizobium m
30	159.5	6.8	211	16	Q8YK44	Q8YK44 arabidena sp
31	156.5	6.7	239	16	Q9KX6	Q9KX6 streptomyce
32	154.5	6.6	225	2	Q8GFG4	Q8GFG4 rhodococcus
33	151	6.4	200	16	Q8GFG4	Q8GFG4 rhizobium 1
34	149.5	6.4	218	2	Q8GFG3	Q8GFG3 rhodococcus
35	149.5	6.4	221	16	Q82Q14	Q82Q14 streptomyce
36	147.5	6.3	226	2	Q8GFK6	Q8GFK6 rhodococcus
37	147	6.3	213	17	Q29059	Q29059 rhodococcus
38	144.5	6.1	226	2	Q9AH05	Q9AH05 rhodococcus
39	143.5	6.1	223	17	Q8YQ9	Q8YQ9 methanopyru
40	142	6.0	212	17	Q29370	Q29370 archaeoglob
41	142	6.0	224	1	Q59661	Q59661 methobact
42	137.5	5.8	199	2	Q9F418	Q9F418 mycobacteri
43	136.5	5.8	217	2	Q83VCI	Q83VCI streptomyce
44	136.5	5.8	217	16	Q9JN78	Q9JN78 streptomyce
45	135.5	5.8	203	2	Q8GFG1	Q8GFG1 rhodococcus

ALIGNMENTS

RESULT 1	ID	Q81UE7	PRELIMINARY;	PRT;	490 AA.
AC	Q81UE7				
DT	01-MAR-2003	(T-EMBLrel.. 23, Created)			
DT	01-MAR-2003	(T-EMBLrel.. 23, Last sequence update)			
DT	01-JUN-2003	(T-EMBLrel.. 24, Last annotation update)			
DE	STAMP1.				
OS	Homo sapiens (human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Prostate;				
RX	MEDLINE=2229309; PubMed=12095985;				
RA	Korkmaz K.S., Blbl C.C., Korkmaz C.G., Loda M., Hager G.L.,				
RA	Saatchioglu F.;				
RT	"Molecular cloning and characterization of STAMP1, a highly prostate				
RT	specific six-trans-membrane protein that is overexpressed in prostate				
RT	cancer."				
RL	J. Biol. Chem. 277:36689-36696(2002).				
DR	EMBL; AY008445; AAC32149.1; ..				
DR	GO; GO:0006118; Electon transport; IEA.				
DR	InterPro; IPR004455; NADPoxred_F420.				
DR	Pfam; PF03807; F420 oxidored; 1.				
DR	SEQUENCE 490 AA; 56027 MW; D1971A8480F7551 CRC64;				

QY	1	MESISMWGSFKSLSETCLP	97.4%;	Score 2290;	DB 4;	Length 490;
QY	1	MESISMWGSFKSLSETCLP	99.8%;	Pred. No. 4.1e-176;		
DB	1	MESISMWGSFKSLSETCLP	0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	61	RNPFASEFPFPHVVDVTHEDALTKNIIFVAIHRREHYTSIMPLRLHLVQKILIDVSNMN				
DB	61	RNPFASEFPFPHVVDVTHEDALTKNIIFVAIHRREHYTSIMPLRLHLVQKILIDVSNMN				
QY	120	61				
DB	120	61				

```

QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVYSAMALQGPDKASROYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVYSAMALQGPDKASROYICSNNIQARQVIE 180
QY 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPVIAITLLSVYLAGLAAAYQLYYGTXYRRPPEMLTWTQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPVIAITLLSVYLAGLAAAYQLYYGTXYRRPPEMLTWTQ 300
QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYLFNNAYQOVHANIENSNWEEVRIEMV 360
DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYLFNNAYQOVHANIENSNWEEVRIEMV 360
QY 361 ISFGIMSLGLSLAATSI PSVSNALNMRSEFSIOSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAATSI PSVSNALNMRSEFSIOSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRYFTPPNPFVLAIVLPSIVIL 444
DB 421 EBYRYFTPPNPFVLAIVLPSIVIL 444

```

RESULT 2

```

Q8NFT2 PRELIMINARY; PRT; 490 AA.
ID Q8NFT2;
AC Q8NFT2;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of prostate 2.
GN STRAP2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Prostate;
RA Forxka K.P., Helentius W.A., Visakorpi T.;
RT "Cloning and characterization of a novel six-transmembrane protein
RT STEAP2, down-regulated in androgen-independent prostate cancer.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF455138; AN04080.1; -
DR Genew; HGNC:17885; STRAP2.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0006118; P: electron transport; IEA.
DR InterPro: IPR004455; NADPoxred_F420.
DR Pfam: PF03807; F420_oxidored; 1.
KW Transmembrane.
SQ SEQUENCE 490 AA; 55961 MW; F2B9C30CDACCEFF81 CRC64;

```

Query Match 97.2%; Score 2284; DB 4; Length 490;

Best Local Similarity 99.8%; Pred. No. 1.2e-175;

Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MESISMWGPXSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVVIGS 60
DB 1 MESISMWGPXSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVVIGS 60
QY 61 RNKFASEFPFHVDVTHHEDALTKNIIIFVAIHREHYTSLMDRLHLVKGKILIDVSNM 120
DB 61 RNKFASEFPFHVDVTHHEDALTKNIIIFVAIHREHYTSLMDRLHLVKGKILIDVSNM 120
QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVYSAMALQGPDKASROYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVYSAMALQGPDKASROYICSNNIQARQVIE 180
QY 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240

```

```

DB 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPVIAITLLSVYLAGLAAAYQLYYGTXYRRPPEMLTWTQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPVIAITLLSVYLAGLAAAYQLYYGTXYRRPPEMLTWTQ 300
QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYLFNNAYQOVHANIENSNWEEVRIEMV 360
DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYLFNNAYQOVHANIENSNWEEVRIEMV 360
QY 361 ISFGIMSLGLSLAATSI PSVSNALNMRSEFSIOSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAATSI PSVSNALNMRSEFSIOSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRYFTPPNPFVLAIVLPSIVIL 444
DB 421 EBYRYFTPPNPFVLAIVLPSIVIL 444

```

RESULT 3

```

Q8BWB6 PRELIMINARY; PRT; 489 AA.
ID Q8BWB6;
AC Q8BWB6;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Weakly similar to tumor suppressor PHDE.
GN 492158B17Rik.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Head;
RA MEDLINE=2234683; PubMed=12466851;
RA The FANTOM Consortium;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL: AK052981; BAC5230.1; -
DR MGD: MGI:1323301; 492153B817Rik.
DR GO: GO:0006118; P: electron transport; IEA.
DR InterPro: IPR004455; NADPoxred_F420.
DR Pfam: PF03807; F420_oxidored; 1.
SQ SEQUENCE 489 AA; 55760 MW; 98CD63D59DDDF24C CRC64;

```

Query Match 95.3%; Score 2239.5; DB 11; Length 489;

Best Local Similarity 97.3%; Pred. No. 4.8e-172;

Matches 432; Conservative 9; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 MESISMWGPXSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVVIGS 60
DB 1 MESISMWGPXSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVVIGS 59
QY 61 RNKFASEFPFHVDVTHHEDALTKNIIIFVAIHREHYTSLMDRLHLVKGKILIDVSNM 120
DB 61 RNKFASEFPFHVDVTHHEDALTKNIIIFVAIHREHYTSLMDRLHLVKGKILIDVSNM 119
QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVYSAMALQGPDKASROYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVYSAMALQGPDKASROYICSNNIQARQVIE 179
QY 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 240
DB 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVRDVIHPYA 239
QY 241 RNOQSDFYKIPIEIVNKTLPVIAITLLSVYLAGLAAAYQLYYGTXYRRPPEMLTWTQ 300
DB 241 RNOQSDFYKIPIEIVNKTLPVIAITLLSVYLAGLAAAYQLYYGTXYRRPPEMLTWTQ 299
QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYLFNNAYQOVHANIENSNWEEVRIEMV 360
DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYLFNNAYQOVHANIENSNWEEVRIEMV 360

```


QY 424 YREYTPNFVLAIVLPSIVIL 444
 DB 424 YKFLPPTFTLTLLPCVILL 444

RESULT 6

Q8C159 PRELIMINARY; PRT; 488 AA.
 ID Q8C159;
 AC Q8C159;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Similar to RIKEN cDNA 1010001D01 gene.
 GN TSAP6 OR 1010001D01R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH 11;
 RA Strausberg R.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC037435; AA037435.1;
 DR MGD; MGI:1915678; Tsap6.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420_oxidored; 1.
 SQ SEQUENCE 488 AA; 54749 MW; 9A08D99C90C83F4 CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 488;
 Best Local Similarity 56.2%; Pred. No. 1.2e-97;
 Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPENGINGIKDARKVTVGVIGSGDFASLTIRLCGYHVVIGSRNPKF 65
 DB 14 LVSDSGSLAE--VF-----KEAPK--VGILSGDFASLTIRLVGSGFVVGSRNPKR 63
 QY 66 ASEFPFHVVDVTHEDALTKNTIIFVAIHREHYTSLMDLRHLVYKILIDVSNMRIN-- 123
 DB 64 TAGLFPSLAQVTFQBEAVSSPEVIFVAVFREHYTSLCGLADQLACKILVDVSNPEKHL 123
 QY 124 QYPESNAEYTLASLPDLSLVKGFNVVSAMALQGFKDSROYTISNNIQARQVIELAR 183
 DB 124 QHRQNAEYTLASLPACVVKAFNVISAMALQAGRDNRQVLCSDQPEAKRTISEVAR 183
 QY 184 QLNFTPIDGLSSAREIENLPLRLFTLMRGVVAISLATFFLYSVRDIHPYRANQ 243
 DB 184 AMGFPLDGLSASAREVEAIPRLPLSKVPTLLALGLFVCFYTNFIRVDLOPIYRKD 243
 QY 244 QSDFYKPIPIEYNTKLPPIVAITLLSLVYLAGLAAAYQYVGTKYRPPMLWTMLQCRK 303
 DB 244 ENKFKMPLSVVNTLPCVAVYLLSLVYLPVLAALQLRGTQYQRPDMLDHVLQHRK 303
 QY 304 QIGLISFFPAMLAHAYSLCLPFRSRHYDLVLAQVANKSRMLWEEVWRMEIYISL 363
 DB 304 QIGLISFFPAMLAHAYSLCLPFRSRHYDLVLAQVANKSRMLWEEVWRMEIYISL 363
 QY 364 GIMSIGLSLLAVTSIPVSNALNMRSEFFIOSTLGYVALLISTHVLITYGWTRAFEEHN 423
 DB 364 GVALGMSLLAVTSIPVSANSLNMRSEFFVOSTLGFVALLISTHVLITYGWTRAFEEHN 423
 QY 424 YREYTPNFVLAIVLPSIVIL 444
 DB 424 YKFLPPTFTLTLLPCVILL 444

RESULT 7

Q80ZF3 PRELIMINARY; PRT; 526 AA.
 ID Q80ZF3;
 AC Q80ZF3;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Tsap6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSMZ ACC331; TISSUE=Bone marrow;
 RX MEDLINE=22506415; PubMed=12606722;
 RA Passer B.J., Nancy-Porbocis V., Amzallag N., Prieur S., Gans C.,
 RA Roborel de Climens A., Pucci G., Bouvard V., Tuyen M., Sueti L.,
 RA Morchoisne S.P., Cribble V., Leepaghol A., Dausset J., Oren M.,
 RA Amson R., Telerman A.;
 RT "The p53-inducible TSAP6 gene product regulates apoptosis and the cell
 cycle and interacts with Nix and the Myl kinase."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:2284-2289(2003).
 DR EMBL; AY214462; AAC38239.1;
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420_oxidored; 1.
 SQ SEQUENCE 526 AA; 58530 MW; 6306CD717E25200A CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 526;
 Best Local Similarity 56.2%; Pred. No. 1.3e-97;
 Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPENGINGIKDARKVTVGVIGSGDFASLTIRLCGYHVVIGSRNPKF 65
 DB 52 LVSDSGSLAE--VF-----KEAPK--VGILSGDFASLTIRLVGSGFVVGSRNPKR 101
 QY 66 ASEFPFHVVDVTHEDALTKNTIIFVAIHREHYTSLMDLRHLVYKILIDVSNMRIN-- 123
 DB 102 TAGLFPSLAQVTFQBEAVSSPEVIFVAVFREHYTSLCGLADQLACKILVDVSNPEKHL 161
 QY 124 QYPESNAEYTLASLPDLSLVKGFNVVSAMALQGFKDSROYTISNNIQARQVIELAR 183
 DB 162 QHRQNAEYTLASLPACVVKAFNVISAMALQAGRDNRQVLCSDQPEAKRTISEVAR 221
 QY 184 QLNFTPIDGLSSAREIENLPLRLFTLMRGVVAISLATFFLYSVRDIHPYRANQ 243
 DB 222 AMGFPLDGLSASAREVEAIPRLPLSKVPTLLALGLFVCFYTNFIRVDLOPIYRKD 281
 QY 244 QSDFYKPIPIEYNTKLPPIVAITLLSLVYLAGLAAAYQYVGTKYRPPMLWTMLQCRK 303
 DB 282 ENKFKMPLSVVNTLPCVAVYLLSLVYLPVLAALQLRGTQYQRPDMLDHVLQHRK 341
 QY 304 QIGLISFFPAMLAHAYSLCLPFRSRHYDLVLAQVANKSRMLWEEVWRMEIYISL 363
 DB 342 QIGLISFFPAMLAHAYSLCLPFRSRHYDLVLAQVANKSRMLWEEVWRMEIYISL 401
 QY 364 GIMSIGLSLLAVTSIPVSNALNMRSEFFIOSTLGYVALLISTHVLITYGWTRAFEEHN 423
 DB 402 GVALGMSLLAVTSIPVSANSLNMRSEFFVOSTLGFVALLISTHVLITYGWTRAFEEHN 461
 QY 424 YREYTPNFVLAIVLPSIVIL 444
 DB 462 YKFLPPTFTLTLLPCVILL 482

RESULT 8

Q92421 PRELIMINARY; PRT; 514 AA.
 ID Q92421;
 AC Q92421;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE DuduLin 2.
 GN TSAP6 OR 1010001D01R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;
 RN SEQUENCE FROM N.A.
 RA Serru V., Lambdin D., Lencir C., Manivet P., Vaudourdoile M.,
 RA Kelleman O., Loric S.,
 RT "Molecular cloning and expression of mouse dudulin 2,"
 RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY029586; AKS0539.1; -
 DR MGD: MGI:1915678; Tsape.
 DR GO: GO:0006118; P.electron transport; IEA.
 DR InterPro: IPR004455; NADPoxred_F420.
 DR Pfam: PF03807; F420 oxidored; I.
 DR PROSITE: PS00290; IE_MHC; I.
 DR SEQUENCE 514 AA; 57268 MW; 3398866C288AEC0E2 CRC64;
 Query Match 55.6%; Score 1306; DB 11; Length 514;
 Best Local Similarity 56.2%; Pred. No. 8.3e-97;
 Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;
 QY 6 MMSGPKLSSETCLPNGINGIDAKKAVTVGVIGSGDFPKSLTIRIRCGHYVIGSRNPK 65
 Db 14 LVSDGSLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSFVVGSRNPKR 63
 QY 66 ASEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRLHVLGKILIDVSNMRLN-- 123
 Db 64 TAGLFPSLAQVTFQEEAVSSPEVIFVAVFRHYSLSGLADQLAGKILVDVSNPTKEHL 123
 QY 124 QYPSNAEYLAFLPDSLYKGFNVSAALQLGPKDASROYIGSNNIQARQCYIELAR 183
 Db 124 QHROSNEVYLAFLPACTVVAFFVVISAMALQAVPRGDRQVLLCGNDKAKRVMEMAR 183
 QY 184 QLNFPIDGLSSAREIENPLRLFTLRGVPVVAISLAEFFLYSFVRDVIHPYARNQ 243
 Db 184 AMGFLPDMGSLASREVEALPLRLBSGKPTLLALGLFVCFYTFIDVLOPIRKD 243
 QY 244 QSDFYKIPLEIVNKLPIVAITLLSLVYLAGLAAVQLYGKIRPPEMLETWLQCRK 303
 Db 244 ENKFKMPLSVNTTLPCVAYVLLSLVYLPGLAAALQRRGTXYORFPDMLDHWLQHRK 303
 QY 304 QLGILSFFPAMVAVYSLCLPMRSEERYFLNMAVQOVHANISNMEEEVRIEYISF 363
 Db 304 QIGLSIFFPAMVAVYSLCLPMRSHRYDLVNLAKVQLANKSLNMBEEMRIEYISL 363
 QY 364 GIMSGLSLAVTSIPSVSNALNMRSEFISTGLGVALLISTFHYLYGKMAFEERY 423
 Db 364 GVALGMSLAVTSIPSIANSINMRSEFVOSTLGFVALLISTHLYTGMTRAFEDN 423
 QY 424 YRFYTPNFVALVLPISIVIL 444
 Db 424 YKFYLPPTFTLLPVCVIL 444
 RESULT 9
 Q86SF6 PRELIMINARY; PRT; 488 AA.
 ID Q86SF6
 AC Q86SF6
 DT 01-UN-2003 (Tremblrel. 24, Created)
 DT 01-UN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Similar to dudulin 2 (Tsape).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN SEQUENCE FROM N.A.
 RA Straussberg R.;
 RA Straussberg R.;
 RT Submitted (Jan-2003) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RX MEDLINE:22506415; PubMed:12606722;
 RA Passer B.U., Nancy-Porteboom V., Amzallag N., Prieur S., Caris C.,
 RA Roborel de Climens A., Finucci G., Bouvard V., Tuyen M., Steini L.,

Morchoisne S.P., Cribble V., Leespagnol A., Daueset J., Oren M.,
 RA Amson R., Teitelman A.;
 RT "The p53-inducible TSAB gene product regulates apoptosis and the cell
 RT cycle and interacts with Nix and the Myt1 kinase,"
 RT Proc. Natl. Acad. Sci. U.S.A. 100:2284-2289(2003).
 DR EMBL: BC042150; AAH42150.1; -
 DR EMBL: AY214461; AAC38238.1; -
 DR GO: GO:0006118; P.electron transport; IEA.
 DR InterPro: IPR003006; IE_MHC.
 DR InterPro: IPR004455; NADPoxred_F420.
 DR Pfam: PF03807; F420 oxidored; I.
 DR PROSITE: PS00290; IE_MHC; I.
 DR SEQUENCE 488 AA; 54600 MW; C69EB0D0430F9BBB CRC64;
 Query Match 54.1%; Score 1272; DB 4; Length 488;
 Best Local Similarity 54.5%; Pred. No. 4.3e-94;
 Matches 242; Conservative 82; Mismatches 108; Indels 12; Gaps 4;
 QY 3 SISMMGPKLSSETCLPNGINGIDAKKAVTVGVIGSGDFPKSLTIRIRCGHYVIGSRN 62
 Db 11 SHLVDSOSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGSFVVGSRN 60
 QY 63 PKFASEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRLHVLGKILIDVSN--NM 120
 Db 61 PKTRALFPSSAQVTFQEEAVSSPEVIFVAVFRHYSLSGLADQLAGKILVDVSNPTKEH 120
 QY 121 RINQYPSNAEYLAFLPDSLYKGFNVSAALQLGPKDASROYIGSNNIQARQCYIELAR 180
 Db 121 EHLQRESNAEYLAFLPCTVVAFFVVISAMTLQAGRDGRQVLLCGNDKAKRVMEMAR 180
 QY 181 LARQNFIPIDGLSSAREIENPLRLFTLRGVPVVAISLAEFFLYSFVRDVIHPYARNQ 240
 Db 181 MALAMGFLPDMGSLASREVEALPLRLBSGKPTLLALGLFVCFYTFIDVLOPIRKD 240
 QY 241 RNOQSDFYKIPLEIVNKLPIVAITLLSLVYLAGLAAVQLYGKIRPPEMLETWLQCRK 300
 Db 241 QSNQNFPLPVSNTTLPCVAYVLLSLVYLPGLAAALQRRGTXYORFPDMLDHWLQ 300
 QY 301 CRRQGLSFFPAMVAVYSLCLPMRSEERYFLNMAVQOVHANISNMEEEVRIEYISF 360
 Db 301 HRKQGLSFFPAMVAVYSLCLPMRSHRYDLVNLAKVQLANKSLNMBEEMRIEYISL 360
 QY 361 ISFGMSGLSLAVTSIPSVSNALNMRSEFISTGLGVALLISTFHYLYGKMAFEERY 420
 Db 361 LSGVALLAGTLLSLAVTSIPSIANSINMRSEFVOSTLGFVALLISTHLYTGMTRAF 420
 QY 421 EEFYTPNFVALVLPISIVIL 444
 Db 421 ESRYKFLPPTFTLLPVCVIL 444
 RESULT 10
 Q72389 PRELIMINARY; PRT; 498 AA.
 ID Q72389
 AC Q72389
 DT 01-OCT-2003 (Tremblrel. 25, Created)
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Hypothetical protein DKFZ668H07150.
 GN DKFZ668H07150.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN SEQUENCE FROM N.A.
 RA Tissue=human colon endothe primary cell culture;
 RA Bloecker H., Boecker M., Mewes H.W., Well B., Amid C., Oeanger A.,
 RA Fobo G., Han M., Wiemann S.;
 RT Submitted (Jan-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BK538047; CAD97986.1; -
 KW Hypothetical protein.
 DR SEQUENCE 498 AA; 55593 MW; 69DD2D5A00C2D6B CRC64;


```

Db      61 PKRIARLPFAAQTPOEBAVSSPEVIFVAVFRHHYSSLSLSQLAGKILVDVSNPTQ 120
Qy      121 RINQYPSNAEYLAFLPDSILYKGFNVASAMALQCPKASROYICSNNTIQAQCVIE 180
Db      121 EHLQRESNAEYLAFLPCTCTVVAFAVVISAMTLOAGPRDGNROVPICGQPEAKVASE 180
Qy      181 LARQNFIPIDGLSSAREIENLPLRFLTMGRPVVAISLAFELFYSFVRDVHPYA 240
Db      181 MALMGFMPVDMGSLASWEVEAMPRLPLPMKPTLIALGLFCFAVNFVRDVLQPY 240
Qy      241 RMOQSDPKPIEIVNKTLPVATLTLVYLAGLAAAYQLYGTYKRRPPEMLTQ 300
Db      241 QESQNKFFKLPSVSVNTLPCVAYVTLSTVYLPGLAAALQRRGTXYORFPDWLDHMQ 300
Qy      301 CRKQGLISFFPAAVAVASLCLPMRSEERYLFNMAVQOVHANIENSNNEEYVRIEM 360
Db      301 HRKQIGLISFFCAALHALVSCFLPRARHYDVNLAVKQVLANKSHLM-VEEVRMEIY 359
Qy      361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIQSTLGYVALLISTFHVLIYGMKRAFE 420
Db      360 LSLGLALGTLSLAVTSIPSIANSINMRSEFVQSSIGFVALVSTLHTLYGWTBAFE 419
Qy      421 ERYRFPYPPNFVALVPSIYL 444
Db      420 ESRYKFFLPFTLTLLVPCVIL 443

```

RESULT 13

```

ID      Q8TF03      PRELIMINARY;      PRT;      487 AA.
AC      Q8TF03;
DT      01-JUN-2002 (Tremblrel. 21, Created)
DT      01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT      01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE      Tumor suppressor p53.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Steiner M.S., Allay J.A., Wang C.;
RT      "A novel prostate-derived tumor suppressor";
RL      Submitted (Feb-2000) to the EMBL/Genbank/DBJ databases.
DR      EMBL; AF238864; AAL78206.1; -.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR003006; IG_MHC.
DR      InterPro; IPR004455; NADPoxred_F420.
DR      Pfam; PF03807; F420 oxidored; 1.
DR      PROSITE; PS00290; IG_MHC; 1.
SQ      SEQUENCE 487 AA; 54402 MW; AFF160535906F68 CRC64;

```

Query Match 53.0%; Score 1245.5; DB 4; Length 487;
 Best Local Similarity 53.8%; Pred. No. 5.8e-92;
 Matches 239; Conservative 83; Mismatches 109; Indels 13; Gaps 5;

```

Qy      3 SISMGSPKSLSEFTCLPNGINGIKARXYTVIGSGDPAKSLTIRLCGHVYIGSRN 62
Db      11 SLHLVDSDSLAK--VPD-----EAPKYS--ILSGDPARSLATRLVSGFKYVGSRN 60
Qy      63 PKFASEFPFHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDIRHLVKGKILIDVSN--NM 120
Db      61 PRTARLPFAAQTPOEBAVSSPEVIFVAVFRHHYSSLSLSQLAGKILVDVSNPTQ 120
Qy      121 RINQYPSNAEYLAFLPDSILYKGFNVASAMALQCPKASROYICSNNTIQAQCVIE 180
Db      121 EHLQRESNAEYLAFLPCTCTVVAFAVVISAMTLOAGPRDGNROVPICGQPEAKVASE 180
Qy      181 LARQNFIPIDGLSSAREIENLPLRFLTMGRPVVAISLAFELFYSFVRDVHPYA 240
Db      181 MALMGFMPVDMGSLASWEVEAMPRLPLPMKPTLIALGLFCFAVNFVRDVLQPY 240

```

```

Qy      241 RMOQSDPKPIEIVNKTLPVATLTLVYLAGLAAAYQLYGTYKRRPPEMLTQ 300
Db      241 QESQNKFFKLPSVSVNTLPCVAYVTLSTVYLPGLAAALQRRGTXYORFPDWLDHMQ 300
Qy      301 CRKQGLISFFPAAVAVASLCLPMRSEERYLFNMAVQOVHANIENSNNEEYVRIEM 360
Db      301 HRKQIGLISFFCAALHALVSCFLPRARHYDVNLAVKQVLANKSHLM-VEEVRMEIY 359
Qy      361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIQSTLGYVALLISTFHVLIYGMKRAFE 420
Db      360 LSLGLALGTLSLAVTSIPSIANSINMRSEFVQSSIGFVALVSTLHTLYGWTBAFE 419
Qy      421 ERYRFPYPPNFVALVPSIYL 444
Db      420 ESRYKFFLPFTLTLLVPCVIL 443

```

RESULT 14

```

ID      Q8TDP3      PRELIMINARY;      PRT;      456 AA.
AC      Q8TDP3;
DT      01-JUN-2002 (Tremblrel. 21, Created)
DT      01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT      01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE      p53.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Wang C., Allay J.A., Steiner M.S.;
RT      "Second human member of p53 family, Human p53.2";
RL      Submitted (May-2000) to the EMBL/Genbank/DBJ databases.
DR      EMBL; AF262322; AAO08128.1; -.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR003006; IG_MHC.
DR      InterPro; IPR004455; NADPoxred_F420.
DR      Pfam; PF03807; F420 oxidored; 1.
DR      PROSITE; PS00290; IG_MHC; 1.
SQ      SEQUENCE 456 AA; 50430 MW; C5F7C700B525251E CRC64;

```

Query Match 45.7%; Score 1074.5; DB 4; Length 456;
 Best Local Similarity 52.9%; Pred. No. 3.2e-78;
 Matches 209; Conservative 75; Mismatches 98; Indels 13; Gaps 5;

```

Qy      3 SISMGSPKSLSEFTCLPNGINGIKARXYTVIGSGDPAKSLTIRLCGHVYIGSRN 62
Db      11 SLHLVDSDSLAK--VPD-----EAPKYS--ILSGDPARSLATRLVSGFKYVGSRN 60
Qy      63 PKFASEFPFHVVDVTHHEDALTKTNIIFVAIHREHYTSLMDIRHLVKGKILIDVSN--NM 120
Db      61 PRTARLPFAAQTPOEBAVSSPEVIFVAVFRHHYSSLSLSQLAGKILVDVSNPTQ 120
Qy      121 RINQYPSNAEYLAFLPDSILYKGFNVASAMALQCPKASROYICSNNTIQAQCVIE 180
Db      121 EHLQRESNAEYLAFLPCTCTVVAFAVVISAMTLOAGPRDGNROVPICGQPEAKVASE 180
Qy      181 LARQNFIPIDGLSSAREIENLPLRFLTMGRPVVAISLAFELFYSFVRDVHPYA 240
Db      181 MALMGFMPVDMGSLASWEVEAMPRLPLPMKPTLIALGLFCFAVNFVRDVLQPY 240
Qy      241 RMOQSDPKPIEIVNKTLPVATLTLVYLAGLAAAYQLYGTYKRRPPEMLTQ 300
Db      241 QESQNKFFKLPSVSVNTLPCVAYVTLSTVYLPGLAAALQRRGTXYORFPDWLDHMQ 300
Qy      301 CRKQGLISFFPAAVAVASLCLPMRSEERYLFNMAVQOVHANIENSNNEEYVRIEM 360
Db      301 HRKQIGLISFFCAALHALVSCFLPRARHYDVNLAVKQVLANKSHLM-VEEVRMEIY 359
Qy      361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIQ 395
Db      360 LSLGLALGTLSLAVTSIPSIANSINMRSEFVQ 394

```

RESULT 15

Q923B6 PRELIMINARY; PRT; 470 AA.
 AC Q923B6;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
 DE Similar to hypothetical protein FLJ23153 (Tnf α -induced adipose-related protein).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Aorta, and Vein;
 RX MEDLINE=22354683; PubMed=12466651;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; BC006651; AA006651.1; -;
 DR EMBL; AK040760; BAC00696.1; -;
 DR GO; GO:0005118; Pilelectron transport; IEA.
 DR InterPro; IPR000585; Hemopexin.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420_oxidored; 1.
 DR PROSITE; PS00024; HEMOPEXIN; 1.
 DR SEQUENCE 470 AA; 52994 MW; 6823E7682AC78B19 CRC64;
 SQ

Query Match 45.3%; Score 1064.5; DB 11; Length 470;
 Best Local Similarity 48.5%; Pred. No. 2.1e-77;

Matches 200; Conservative 80; Mismatches 131; Indels 1; Gaps 1;

QY 33 VGVIGSGDFPKSITLRLIRGHHVYIGSNPKFSEFPFHVDVTHEDALTKNIIFVA 92
 DB 22 VCIFFGTGDFGKSLGLKMLCGYSIVGSRNPQ-VSSLPRGAELVSYEEAASKSDIITLA 80
 QY 93 IREHRTSITMDLPHLLVKGILLDVSNMKNINQYPSNNAEYLASTLPDLSLVKGFNVASAM 152
 DB 81 MREHYDLSLELDVLYKGLVDSNNRKINQYPSNNAEYLASTLPDLSLVKGFNVASAM 140
 QY 153 ATOLGPKDASRQYVIGCSNNIQARQYIEHARQINFIPIIDGLSSAREIENLPLRLFTLM 212
 DB 141 ALQSGTLDASRQYFVCGNSKAKQRMVDAITLGLPLDQSLMASEIENYPLQLFPMW 200
 QY 213 RGPVVVAISLATFFELYSFVRDVIHPYANQSDFKYKIPRIYVNTLPIVATLLSLVYL 272
 DB 201 RPFVYLSVLCVFFFYCAIREVIYVYNGKTDATYRLAISIPNRVFPITLILALVYL 260
 QY 273 AGLLAAAYOLYVGTGKRRFPFMTLQCRKQGLISFFFAWVAVASLCPMRSERYL 332
 DB 261 PGLIAAILQYRKTFRFPFNMLDHMLCKRKGVALGAFALHVIYTLVIRYVVRWR 320
 QY 333 FLNMAVQCHANIENSNNEEVRIRIEMYSISFGIMSGLLSLAVTSIPSVSNALNMRFS 392
 DB 321 LRNAVITTOALTNDOSPITSYAMINDSYALGILGFLLGLITSLPSVSNVNMWREPR 380
 QY 393 FIOSTLGVALLISTFHVLIYGMKRAFEERYRFRYTPPNFVALVLPISVIL 444
 DB 381 FVQSKGLYLVLCITATLIVYGGKRFSLPILMSLPSAVITLAVIPCAVLV 432

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:33:12 ; Search time 17 Seconds
(without alignments)
1390.579 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMWGPSEKISELCPN.....ALVLPISVLDLQLCRYPD 454

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	717	30.5	339	1	STEAL_HUMAN
2	133	5.7	223	1	Q58896 methanococ
3	120.5	5.1	232	1	Q26350 methanobact
4	111.5	4.7	695	1	P32212 macaca fasc
5	109.5	4.7	206	1	Q09097 pasteuralla
6	109	4.6	693	1	P78763 gallus gall
7	109	4.6	712	1	Q12473 saccharomyc
8	107.5	4.6	629	1	Q12343 saccharomyc
9	104.5	4.6	695	1	P23945 homo sapien
10	104.5	4.4	695	1	P48059 sus scrofa
11	104	4.4	396	1	P58529 escherichia
12	104	4.4	396	1	P31122 escherichia
13	103	4.4	465	1	O61983 mus musculu
14	102.5	4.4	391	1	O84K66 nephroselimi
15	102	4.3	220	1	O84V59 bruceella me
16	102	4.3	220	1	O84V59 bruceella me
17	100	4.3	476	1	P05512 saccharomyc
18	99.5	4.2	452	1	O79421 brachiolesto
19	99	4.2	365	1	O57443 homo sapien
20	98	4.2	3411	1	P03314 y genome po
21	98	4.2	3411	1	P03314 y genome po
22	97.5	4.1	268	1	O61364 aquilex deo
23	97.5	4.1	695	1	P33376 bos taurus
24	97	4.1	218	1	O84B19 xanthomonas
25	97	4.1	692	1	P20395 rattus norv
26	96.5	4.1	686	1	P33791 saccharomyc
27	96	4.1	261	1	P77409 escherichia
28	95	4.0	202	1	P43409 pseudomonas
29	95	4.0	393	1	P56753 arabidopsis
30	95	4.0	693	1	Q58135 sprinacia oi
31	94.5	4.0	992	1	P33378 mus musculu
32	94.5	4.0	492	1	P25014 cyanophora
33	94.5	4.0	695	1	P35379 ovine aries

34	93.5	4.0	365	1	SPR1_MOUSE
35	93.5	4.0	452	1	NU4M_BRAF1
36	93.5	4.0	499	1	Q15770 plasmodium
37	93.5	4.0	530	1	AAAI_MOUSE
38	93.5	4.0	1233	1	VIA1_BROVD
39	93	4.0	388	1	HMC3_DESVH
40	93	4.0	459	1	NU4M_BOVIN
41	93	4.0	521	1	YT25_CAEEL
42	92.5	3.9	490	1	GTRI_CHICK
43	92.5	3.9	500	1	NU4C_OENHO
44	92.5	3.9	883	1	YHL6_YEAST
45	92.5	3.9	1769	1	YRK9_YEAST

ALIGNMENTS

RESULT 1
ID STEAL_HUMAN STANDARD; PRT; 339 AA.
AC Q9UEB8; O95034;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Six transmembrane epithelial antigen of prostate.
GN STEAP OR STEAPL
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP MEDLINE=20056277; PubMed=10588738;
RA Hubert R.S., Vivanco I., Chen B., Raategar S., Leong K.,
RA Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Rattano A.B.,
RA Jakobovits A., Saffian D.C., Afar D.E.H.;
RT "STEAP: a prostate-specific cell-surface antigen highly expressed in
human prostate tumors."
RL Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).
RL [2]
RP SEQUENCE FROM N.A.
RA Abu-Threiden U., Stoneking T., Langston Y., Maupin R.;
RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP TISSUE=Skid;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Burow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Braletstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshvuk S., Carlini P., Mullaly S.J.,
RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultx S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hulton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywicki M.I., Skalski U., Smallus D.E.,
RA Scheraga A., Schein U.E., Jones S.J.M., Maita M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC - SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC - TISSUE SPECIFICITY: Highly expressed in prostate tumors.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC reduction of NADP(+) with F420H(2). Probably couples the NADP-
CC dependent oxidation of the alcohol to the aldehyde with the F420-
CC dependent reduction of Co(2) to methane (anaerobic function) (By
CC similarity).

CC -1 CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced
CC coenzyme F420.

CC -1 SIMILARITY: TO M.ORGANOPIILUM F420-DEPENDENT NADP REDUCTASE.

CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on the
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
CC or send an email to license@ebi.ac.uk).

CC -----

DR EMBL; U67591; AAB9514.1; -
DR PIR; D64487; D64487.
DR TIGR; MJ1501; -
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; 1.
DR TIGRFAMs; TIGR00301; TIGR00301; 1.
KW Hypothetical protein; Oxidoreductase; NADP. Complete proteome.
SQ SEQUENCE 223 AA; 24068 MW; 2370BDBE5BD39 CRC64;

Query Match 5.7%; Score 133; DB 1; Length 223;
Best Local Similarity 21.9%; Pred. No. 0.003;
Matches 46; Conservative 152; Mismatches 78; Indels 34; Gaps 7;

QY 37 GSGPFASTIRLETRCGVHVIVSGRNKFASEFPFPHVVDV-----THEDALT 84
DB 8 GTGGQGGGLAIRLAK-NKKIISGRKKEKAEAKKAKKEIKKGIGLEADITGLNKKDAK 66
QY 85 KTNIIFFVAIRHER-YTSLMDLRHLVJGKILDV-----SNMRINQYF-SNAEYL 133
DB 67 EGDVYILSLPEYTLSTIKQLKEELKGIIVSIGVPLATAGDPKTRLLFPDPSVAMV 126
QY 134 ASLPEPDLIVGVFVWSAMALQLDPKASRQVYICSNNGAQOQYELARQLNFI-PIDL 192
DB 127 QNVLKESKVSVAFQVNCVAVTEDDNPVDCDIIVGNDSEAKKVVIDLANDIDGRAIDC 186

QY 193 GSLSGAREIENLPLRLFTLWEGPVVAISL 222
DB 187 GNLEKSRILEIAIT-----PLLIGIML 207

RESULT 3
F420_METH
ID F420_METH STANDARD; PRT; 232 AA.
AC 026350;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Purative F420-dependent NADP reductase (EC 1.-.-.-).
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1] -
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum M., Pothier B., Qiu D.,
RA Spadofora R., Viciere R., Wang Y., Wlezbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Saker H., Patwell D., Pradnakar S.,
RA McDougall S., Shiner G., Goyal A., Pletkovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltam: functional analysis and comparative genomics";
RL J. Bacteriol 179:7135-7155(1997)

CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.
CC	-1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC	-1- FMH/LSH/TSH subfamily.
CC	-1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as their content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; X74454; CA852463.1; .
DR	PIR; JN08988; JN0898.
DR	HSSP; P23945; IXUN.
DR	InterPro; IPR000276; GPCR_Rhodopsn.
DR	InterPro; IPR001611; LRR.
DR	InterPro; IPR000372; LRR_Nterm.
DR	Pfam; PF00001; 7tm_1; 1.
DR	Pfam; PF00560; LRR_3.
DR	Pfam; PF01462; LRRNT_1.
DR	PRINTS; PRO0237; GPCRHODOPS.
DR	SMART; SMO0013; LRRNT; 1.
DR	PROSITE; PS00237; G_PROTEIN_RECCE_Ft_1; 1.
DR	PROSITE; PS02623; G_PROTEIN_RECCE_Ft_2; 1.
KW	G-Protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW	Phosphorylation; Repeat; Leucine-rich repeat.
KV	SIGNAL 1 17 POTENTIAL.
FT	CHAIN 18 695 POLICE STIMULATING HORMONE RECEPTOR.
FT	DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM 367 387 1 (POTENTIAL).
FT	DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM 399 421 2 (POTENTIAL).
FT	DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM 444 465 3 (POTENTIAL).
FT	DOMAIN 466 485 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM 486 508 4 (POTENTIAL).
FT	DOMAIN 509 528 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM 529 550 5 (POTENTIAL).
FT	DOMAIN 551 573 CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM 574 597 6 (POTENTIAL).
FT	DOMAIN 598 608 EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM 609 630 7 (POTENTIAL).
FT	DOMAIN 631 695 CYTOPLASMIC (POTENTIAL).
FT	REPEAT 44 68 LRR 1.
FT	REPEAT 69 93 LRR 2.
FT	REPEAT 119 143 LRR 3.
FT	REPEAT 170 192 LRR 4.
FT	REPEAT 193 216 LRR 5.
FT	REPEAT 218 240 LRR 6.
FT	DISULFID 442 517 BY SIMILARITY.
FT	CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE 695 AA; 78343 MW; D060A237329B5250 CnC64;
Query Match	4.7%; Score 111.5; DB 1; Length 695;
Best Local Similarity	17.0%; Pred. No. 0.56;
Matches 84; Conservativity 75; Mismatches 151; Indels 183; Gaps 16	
D5	HEDALTKTNIFVAIHREHYSLMDRLHLVG-----KLLDIVSNMR 121
D6	98 HEIEIKANNL-LINPEAFQNLNKLILISNGIKSLPVNHKHSFOAKLDIIONDN 156
D7	122 INQVESNAEYLAFLPDGSLIVKGPNVVSAMALOLGPYDASROYLYIGSNNTIQAQQVLEL 181
D8	157 IH-----TIENNSFWGLSFESVIWL-----MNGQOEIHNCX----- 189
D9	182 AROLNFIPIDLGSLSAREINLPLRLTTLRGPPVVAVISLATFFVYSFYRDVTHPYAR 241
D10	190 ---ENGQLDLNLINSNNNNLEELNDVDVHGASGVPIIDISTRHSLPSYGLMKLKRA 246

```

QY 242 NQSDPFKIPETIVNKTPIVATILLSIVY----- 271
DB 247 RSTINLKLTP-----SLEKVALMEASITVPSHCCAFANRRQISELHPICNKSILRQEV 301
QY 272 -----LAGLLAAAYQLYVGYKRRF-----PPLMETWLOCRKQLG 306
DB 302 DYMOTQRQSSSLMEDNNESSYSRGFDWYAFEDLQNEVDVYTCSPKPAFNPCCEDILG 361
QY 307 -----LISFPFAMVVA-----YSLCPMRSRRYFLMAYQOVHANIENS 348
DB 362 YNIRVLWFTSIALTGNIIIVLTLTSQYKLYP-----RFLMCLAFAD----- 408
QY 349 WNEEVRNIEMVYISFGIMSLGLSLAATSIPIVSNALNMR-----FSFIQSTLG 399
DB 409 -----LCIGIVULLIASVDITKQYHNVAIDWQTGACDAGGFIVASELS 456
QY 400 -----YVALLISTFH-----VLIGKRAPEEERYRFPYPPNV-LAL 436
DB 457 VYTLTAITLERTHTHAMQLDCKVYHRAASVVMGWIFAFAPALPFIQISSYMKVSI 516
QY 437 VLPSVILLDLQ 449
DB 517 CLPMDIDSPLSQ 529

RESULT 5
Y538_PASMU STANDARD; PRT; 206 AA.
AC 08CNS7;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Hypothetical protein PM0538.
GN PM0538.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida pm70."
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: Belongs to the UPF0191 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AEO06089; AAK02622.1; -
CC DR HAMAP: MF_01207; -1
CC DR InterPro: IPR007916; UPF0191.
CC DR Pfam: PF05252; UPF0191; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 7 26 POTENTIAL.
CC FT TRANSMEM 41 63 POTENTIAL.
CC FT TRANSMEM 76 98 POTENTIAL.
CC FT TRANSMEM 113 135 POTENTIAL.
CC FT TRANSMEM 142 164 POTENTIAL.
CC FT TRANSMEM 169 191 POTENTIAL.
CC ST TRANSMEM 206 AA; 23653 MW; 2F27772SD7P11194 CRC64;
SQ SEQUENCE

Query Match 4.7%; Score 109.5; DB 1; Length 206;
Best local similarity 23.8%; Pred. No. 0.18;
Matches 56; Conservative 44; Mismatches 72; Indels 63; Gaps 13;

```

```

QY 228 LYSFVADVH-----PYA-----RNQSDPFKIPETIVNKTPIVATILLSIVYAGL 275
DB 1 MSLFRILTHHCCGAPVAMLMVLLSGDESLGADPIEIGHFGFSLTILLIMFILD- 59
QY 276 LAAAYQLYVGYKRRFPPLMETWLOCRKQLGLSFPFAMVVAVSLCPMRSRRYFLFN 335
DB 60 -----KVFYLLK-----QPOLQV-----LRRLQIMWFFVYLVH-YA-----YLALE 96
QY 336 MAYQOVHANIENSWNEEVRNIEMVYISFGIMSLGLSLAATSIPIVSNALNMRFSFIQ 395
DB 97 LGVD-----RSLFQELVNR--GYLITGATAFLITLMAISSYLR--LKGKWFYVL 146
QY 396 STLGVALLISTFHVLIGWKRAPEEERYRFPYPPNV-LALVPSVILLDLQ 450
DB 147 HOLGYVALLLGAH--YVW-----SVKNVTFSSWL--YLLISITILC 184

RESULT 6
FSHR_CHICK
ID FSHR_CHICK STANDARD; PRT; 693 AA.
AC P79763; Q90719;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
DE receptor).
GN FSHR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=97473503; PubMed=9332357;
RA Makabayashi N., Suzuki A., Hoshino H., Nishimori K., Mizuno S.;
RT "The cDNA cloning and transient expression of a chicken gene encoding
RT a follicle-stimulating hormone receptor."
RT Gene 197:121-127(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97057887; PubMed=8902217;
RA You S., Bridgham J.T., Foster D.N., Johnson A.L.;
RT "Characterization of the chicken follicle-stimulating hormone
RT receptor (cFSH-R) complementary deoxyribonucleic acid, and expression
RT of cFSH-R messenger ribonucleic acid in the ovary."
RT Biol. Reprod. 55:1055-1062(1996).
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
CC of this receptor is mediated by G proteins which activate
CC adenylate cyclase.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -1- FSH/LSH/TSH subfamily.
CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D87871; BA13487.1; -
CC DR EMBL: U51097; AAC60030.1; -
CC DR HSSP: P23945; 1XUN.
CC DR InterPro: IPR000276; GPCR_Rhodopn.
CC DR InterPro: IPR001611; LRR_Nterm.
CC DR InterPro: IPR000372; LRR_Nterm.
CC Pfam: PF00001; 7tm_1; 1.

```

PFam: P00560: LRR: 2.
 DR PRINTS: PRO0237: GPCRHOODPSN.
 DR SMART: SM0013: LRRNT; 1.
 DR PROSITE: PS00237: G-PROTEIN_RECEP_F1.1; 1.
 DR PROSITE: PS02662: G-PROTEIN_RECEP_F1.2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 KM SIGNAL
 FT CHAIN 1 17
 FT DOMAIN 18 693
 FT TRANSSEM 367 387
 FT TRANSSEM 388 398
 FT TRANSSEM 399 421
 FT TRANSSEM 422 443
 FT TRANSSEM 444 465
 FT TRANSSEM 466 485
 FT TRANSSEM 486 508
 FT TRANSSEM 509 528
 FT TRANSSEM 529 550
 FT TRANSSEM 551 573
 FT TRANSSEM 574 597
 FT TRANSSEM 598 608
 FT TRANSSEM 609 630
 FT TRANSSEM 631 693
 FT DOMAIN 45 68
 FT REPEAT 69 93
 FT REPEAT 95 118
 FT REPEAT 119 143
 FT REPEAT 148 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 47 47
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 268 268
 FT CONFLICT 4 4
 FT CONFLICT 88 88
 FT CONFLICT 140 140
 FT CONFLICT 174 174
 FT CONFLICT 191 191
 FT CONFLICT 329 329
 FT CONFLICT 693 AA; 78697 MW; 46P986969635A1BEC CRC64;
 SQ SEQUENCE
 Query Match 4.6%; Score 109; DB 1; Length 693;
 Best Local Similarity 17.2%; Pred. No. 0.87; Mismatches 173; Indels 238; Gaps 21;
 Matches 100; Conservative 72; Mismatches 173; Indels 238; Gaps 21;
 QY 18 LKNG-INGIKARQVTVG-----VIGSGDFAKSLTIRIRCGYHVIGSRNPKFASEFFP 71
 DB 61 IPKGFTGHDLEKLEISQNDALTEIEGVFSSL-----PKL----- 97
 QY 72 HVDVDTTHEDALTKNITFVAIHRHYNSLMDRLVWG-----KILI 114
 DB 98 -----HTRIEKANKL-WKIDQDFQHLPSRLTLLSNTGLSPVYVHKVHSQKVL 149
 QY 115 DVSNMRINQYPSNAEYLAFLPDSLIVKGFNVSAVALQGPQDASROYICSNITQA 174
 DB 150 DVQDNHIRTIRBNFMGLSS--ESVILR-----LNKGIOE 184
 QY 175 RQGVETLAQLNFIPIDLGSLSSAREIENLPRLFTLMGPPYVAISLATPEFLYSPRD 234
 DB 185 IKD---HAFNGTCLDELNTSDNNLEKLPKRVQAGVAVLISRTIRSELSHGLE 239
 QY 235 VHPYARNOQSDFYKIPIEIVNKTPIVATITLSIVYAGL----- 275
 DB 240 FIKKLRARSTYKTKLP--DVKKFRLIANFTYPSHCAFNKTKQNTFEYFICSMSPA 297
 QY 276 -----LAAAYQLYGTCK-----RR-----PRLFTWLQCR 302
 DB 298 KODLGQTKRKRRKRSAAEDYISHYGTREGVENEEDYGLCNEVVDVPCSPKDPANPCE 357
 QY 303 KOLG-----LISFFAMVVA-----YSLCLPMRBSERYILNNAVYQOVHAN 344

DB 358 DIMGNVLRVLMFINLAINNTVLLIISQYKLVF-----RFLMGLAFAD----- 408
 QY 345 IENSMNEEBWRIENYISFGIMSLGLSLAVTSIPSYSNLNMRE-----FSPIQ 395
 DB 409 -----LCIGIYLLFIASVDIQTSRYNYVAIDWQTAGCNAAGFTVFA 452
 QY 396 STLGVALLISTYFH-----VLIVGKCAF----- 419
 DB 453 SELSVYTLVLIVLEMHNTITVMOQLNRKRYLRHAYIVNFGMVFVYALPIFGISSYM 512
 QY 420 -----EEYRFFYPPNFVYALVLPISVILDIQLCRY 452
 DB 513 KVSICLPMHIEPFGQAVV--IFLLVNVLAFFVILICICICY 553
 RESULT 7
 ID FRB6_YEAST STANDARD; PRT; 712 AA.
 AC Q12473;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ferric reductase transmembrane component 6 precursor (EC 1.16.1.7)
 DE (Ferric-chelate reductase 6).
 GN FRB6 OR YL051C OR L0593.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8288C / Ab972;
 RX MEDLINE=97313267; PubMed=9159871;
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoorge W.,
 RA Baner V., Brueckner M., Deltis H., Dubois E., Duesterhoelt A.,
 RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
 RA Heuss-Netzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
 RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
 RA Mueller-Auer S., Nentwich U., Obermaier B., Pirvadi E., Pohl T.M.,
 RA Portetelle D., Funelle B., Reckmann S., Rieger M., Rinke M., Rose M.,
 RA Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,
 RA Underwood A.P., Ureastarazu L.A., Vandenbol M., Verhaaselt P.,
 RA Vienderdeels F., Voet M., Volckaert G., Voss H., Wambut R., Medler E.,
 RA Medler H., Zimmermann F.K., Zollner A., Hanl J., Hehlsel J.D.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
 RL Nature 387:87-90(1997).
 CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
 CC -1- COFACTOR: FAD (Probable).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the FRE / CYB5 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use. By non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: Z47973; CAAB8006.1; -
 CC EMBL: Z73156; CA97503.1; -
 CC PIR: S50969; S50969.
 CC GERMOnline: 142046; -
 CC SGD: S0003974; FRB6.
 CC InterPro: IPR002916; Ferric_reduct.
 CC Pfam: PFO1794; Ferric_reduct. 1.
 KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;
 KW FAD; NAD; Glycoprotein; signal; Multigene family.
 FT SIGNAL 1 17
 FT CHAIN 18 712
 FT NP_BIND 493 499
 FT FAD (Potential).
 FT FERRIC REDUCTASE TRANSMEMBRANE COMPONENT

OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A., AND VARIANTS ALA-307 AND SER-680.
 RC TISSUE=ovary;
 RA MEDLINE=91222171; PubMed=1709010;
 RX Mingshish T., Nakamura K., Takakura Y., Ithaki Y., Igarashi M.;
 RT "Cloning and sequencing of human FSH receptor cDNA.";
 RL Biochem. Biophys. Res. Commun. 175:1125-1130(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA MEDLINE=93246012; PubMed=1301382;
 RX Kelton C.A., Cheng S.V., Nugent N.P., Schweickhardt R.L.,
 RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzaja J.B., Luchette C.A.,
 RA Chappel S.C.;
 RT "The cloning of the human follicle stimulating hormone receptor and
 its expression in COS-7, CHO, and Y-1 cells.";
 RL Mol. Cell. Endocrinol. 89:141-151(1992).
 [3]
 RP SEQUENCE FROM N.A.
 RA Tilly L.T., Aihara T., Nishimori K., Dai X.-C., Billig H.,
 RA Kowalski K.I., Perlas E.A., Hsueh A.J.;
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE OF 1-342 FROM N.A.
 RC TISSUE=Testis;
 RA MEDLINE=93075197; PubMed=1359889;
 RX Gromoll U., Gudermann T., Nieschlag E.;
 RT "Molecular cloning of a truncated isoform of the human follicle
 stimulating hormone receptor.";
 RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).
 [5]
 RP SEQUENCE OF 1-51 FROM N.A.
 RA MEDLINE=95011044; PubMed=7926278;
 RX Gromoll U., Dankebar B., Gudermann T.;
 RT "Characterization of the 5' flanking region of the human follicle-
 stimulating hormone receptor gene.";
 RL Mol. Cell. Endocrinol. 102:93-102(1994).
 [6]
 RP SEQUENCE OF 266-695 FROM N.A., AND VARIANT SER-680.
 RA MEDLINE=95000244; PubMed=7916967;
 RX Gromoll U., Ried T., Holzgrave-Grez H., Nieschlag E., Gudermann T.;
 RT "Localisation of the human FSH receptor to chromosome 2p21 using a
 genomic probe comprising exon 10.";
 RL J. Mol. Endocrinol. 12:265-271(1994).
 [7]
 RP 3D-STRUCTURE MODELING OF 49-228.
 RA MEDLINE=96363672; PubMed=8747461;
 RX Jiang X., Dreano M., Buckler D.R., Cheng S., Yehier A., Wu H.,
 RA Hendrickson W.A., el Tayar N.;
 RT "Structural predictions for the ligand-binding region of glycoprotein
 hormone receptors and the nature of hormone-receptor interactions.";
 RL Structure 3:1341-1353(1995).
 [8]
 RP VARIANTS ALA-307; ARG-524 AND SER-680.
 RA MEDLINE=99318093; PubMed=10391209;
 RX Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Pettit N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 [9]
 RP ERRTUM.
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Pettit N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 [10]
 RP VARIANTS ALA-307 AND SER-680.
 RA MEDLINE=22054685; PubMed=12059813;

RA Asatiani K., Gromoll U., Eckardstein S.V., Zitzmann M., Nieschlag E.,
 RA Simoni M.;
 RT "Distribution and function of FSH receptor genetic variants in normal
 men.";
 RL Andrologia 34:172-176(2002).
 [11]
 RP VARIANT OHSS IIR-449.
 RX MEDLINE=22812037; PubMed=12930927;
 RA Vasseur C., Rodien P., Beau I., Desroches A., Gerard C.,
 RA de Poncheville L., Chaplot S., Savagner F., Croze A., Mathieu E.,
 RA Lahlou N., Descamps P., Misrahi M.;
 RT "A chorionic gonadotropin-sensitive mutation in the
 follicle-stimulating hormone receptor as a cause of familial
 gestational spontaneous ovarian hyperstimulation syndrome.";
 RL New Engl. J. Med. 349:753-759(2003).
 [12]
 RP VARIANT OHSS ASN-567.
 RX MEDLINE=22812037; PubMed=12930928;
 RA Smith G., Olatunbosun O., Delbaere A., Pierson R., Vassarot G.,
 RA Costagliola S.;
 RT "Ovarian hyperstimulation syndrome due to a mutation in the
 follicle-stimulating hormone receptor.";
 RL New Engl. J. Med. 349:760-766(2003).
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
 of this receptor is mediated by G proteins which activate
 adenylate cyclase.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P23945-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P23945-2; Sequence=VSP_001953; VSP_001954;
 CC -1- TISSUE SPECIFICITY: Sertoli cells and ovarian granulosa cells.
 CC -1- DISEASE: Defects in FSHR are a cause of ovarian hyperstimulation
 syndrome (OHS) (MIM:608115). OHS is a disorder which occurs
 either spontaneously or most often as an iatrogenic complication
 of ovarian stimulation treatments for in vitro fertilization. The
 clinical manifestations vary from abdominal distention and
 discomfort to potentially life-threatening, massive ovarian
 enlargement and capillary leak with fluid sequestration.
 CC Pathologic features of this syndrome include the presence of
 multiple serous and hemorrhagic follicular cysts lined by
 hyperplastic cells, a condition called hyperreactio lutealis.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC FSH/LSH/TSH subfamily.
 CC -1- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL, M65085; AA52477.1; -
 CC EMBL, S59900; AA826480.1; -
 CC EMBL, M95489; AA52478.1; -
 CC EMBL, X68044; CAA48179.1; -
 CC EMBL, S73199; AAB32071.1; -
 CC EMBL, S73526; AAB32225.1; -
 CC PIR, I57661; ORHUT.
 CC PDB, 1XUN; 15-MAY-97.
 CC GeneW, HGNC:3969; FSHR.
 CC MIM, 136435; -
 CC MIM, 608115; -
 CC GO, GO:0016021; C:integral to membrane; TAS.
 CC GO, GO:0005886; C:plasma membrane; TAS.
 CC GO, GO:0004963; F:follicle stimulating hormone receptor activity; TAS.
 CC GO, GO:0007292; P:female gamete development; TAS.
 CC GO, GO:0005855; P:female gamete development; TAS.
 CC GO, GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.

```

DR GO:0007283; P:spermatogenesis; TAS.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR_3; 1.
DR Pfam: PF01462; LRRNT; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR SMART: SM00013; LRRNT; 1.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing;
KW 3D-structure; Polymorphism; Disease mutation.
FT SIGNAL 1 17
FT CHAIN 1 695
FT DOMAIN 18 366
FT TRANSSEM 367 387
FT DOMAIN 388 398
FT TRANSSEM 399 421
FT DOMAIN 422 443
FT TRANSSEM 444 465
FT DOMAIN 466 508
FT TRANSSEM 509 528
FT TRANSSEM 529 550
FT DOMAIN 551 573
FT TRANSSEM 574 597
FT DOMAIN 598 608
FT TRANSSEM 609 630
FT DOMAIN 631 695
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CARBOHYD 318 318
FT VARSPLIC 224 285

```

```

Query Match 4.6%; Score 107.5; DB 1; Length 695;
Best Local Similarity 17.0%; Pred. No. 1.1;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

```

```

QY 79 HEDALTKNIIFVAHREHYTSLMDRLVVG-----KILIDVSNMR 121
DB 98 HEIRREKANNL-LYNPEAFQNLPLYQYLISNGIKHLPRVHKHSIQKYLIDIONIN 156
QY 122 INQYESNAEYLASLFPDSLIYGFNVSAVALQIGPDASROYTCGNNIQAQVIEL 181
DB 157 IH-----TIERSFVGLSEESVILW-----NKGQIOEHMCA----- 189
QY 182 ARLQNPIDGLSASAREIENLPLRLFLWRGPFVVAISLATFFLSFVDVIHPYAR 241
DB 190 ---FNGTQDELNDNNNLELPDVFHGASGYIILISKRINSLSYGLENKKURA 246
QY 242 NQGSDFYKPIEIVKTL-PIVATLLSLVY-----PMLETWLCQCRKQL 305
DB 247 RSTYMLKKLP-----TEKLVALMEASITYPSCAFANWRQISELHPICNKSLRQE 300
QY 272 -----LAGLLAAAYQVLYGTGYRRF-----PMLETWLCQCRKQL 305
DB 301 VDMYQTRQGRSSLEADNESSYRSGPDWYTFDVLQNEVVDYCSKRPDAFNCEDIM 360
QY 306 G-----LISFFAMQVVA-----YSLCLPMKRSERYFLNNAVQOVHANTEN 347
DB 361 GVALIKVLIWFIISILAITGNIIIVLITTSQYKLTVP-----RFLMCLMAFAD----- 408
QY 348 SNMEEVMIEMWISIGINSIGLSLAVATSIPIVSNALMWRB-----FSPIQSTL 398

```

```

DB 409 -----LCIGYLLILNSVDIHRTSQYHNAIDMQTAGCDAGFFIVFASSEL 455
QY 399 G---YVALLISTFH-----VLITGKKRAPEEERYRFTPPNFV-LA 435
DB 456 SVYTLTATLTERWHTITAMQIDCKVQLRHAASVWVGWIFAFAPAAALFFIGISSYKVS 515
QY 436 LVLPSTVLDDLQ 449
DB 516 ICLPMDIDPSLSQL 529

RESULT 10
FSHR_PIG STANDARD; PRT; 695 AA.
ID FS_HR_PIG
AC P49059; O77514;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictropin receptor).
DE FSHR.
GN Sus scrofa (Pig).
OS Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Cetartiodactyla; Suidae, Sus.
OX Mammalia; Euteleostomi; Cetartiodactyla; Suidae, Sus.
RN NCBI_Taxid=9823;
RP (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=96011644; PubMed=7590277;
RA Remy J.U., Labib-Manssaj Y., Yelie M., Bozon V., Couture L.,
RA Fajot E., Grebert D., Salese R.,
RT "The porcine follicle-stimulating hormone receptor: cDNA cloning, functional expression and chromosomal localization of the gene."
RL Gene 163:257-261(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F.,
RA Barbera A.R.;
RT "Porcine follicle-stimulating hormone receptor.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity of this receptor is mediated by G proteins which activate adenylyl cyclase.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors. FSH/LSH/TSH subfamily.
CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: L31966; AAA66933.1; -.
CC EMBL: AF025377; AAC24981.1; -.
CC HSSP: P23945; 1XUN.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR_3; 1.
DR Pfam: PF01462; LRRNT; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR SMART: SM00013; LRRNT; 1.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.

```

FT	SIGNAL	1	17		POLYIC ACID STIMULATING HORMONE RECEPTOR.
FT	CHAIN	18	695		EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	18	366		1 (POTENTIAL).
FT	TRANSMEM	367	387		CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	388	398		2 (POTENTIAL).
FT	TRANSMEM	399	421		EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	422	443		3 (POTENTIAL).
FT	TRANSMEM	444	465		CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	466	485		4 (POTENTIAL).
FT	TRANSMEM	486	508		EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	509	528		5 (POTENTIAL).
FT	TRANSMEM	529	550		CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	551	573		6 (POTENTIAL).
FT	TRANSMEM	574	597		EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	598	608		7 (POTENTIAL).
FT	TRANSMEM	609	630		CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	631	695		LRR 1.
FT	REPEAT	44	68		LRR 2.
FT	REPEAT	69	93		LRR 3.
FT	REPEAT	119	143		LRR 4.
FT	REPEAT	170	192		LRR 5.
FT	REPEAT	193	216		LRR 6.
FT	REPEAT	218	240		BY SIMILARITY.
FT	DISULFID	442	517		N-LINKED (GLYCNAIC. . .) (POTENTIAL).
FT	CARBOHYD	191	191		N-LINKED (GLYCNAIC. . .) (POTENTIAL).
FT	CARBOHYD	199	199		N-LINKED (GLYCNAIC. . .) (POTENTIAL).
FT	CARBOHYD	293	293		T -> A (IN REF. 1).
FT	CONFLICT	2	2		S -> S (IN REF. 1).
FT	CONFLICT	13	13		V -> A (IN REF. 1).
FT	CONFLICT	60	60		V -> M (IN REF. 1).
FT	CONFLICT	166	166		Q -> H (IN REF. 1).
FT	CONFLICT	215	215		K -> R (IN REF. 1).
FT	CONFLICT	247	247		S -> T (IN REF. 1).
FT	CONFLICT	257	257		D -> N (IN REF. 1).
FT	CONFLICT	334	334		E -> K (IN REF. 1).
FT	CONFLICT	349	349		T -> A (IN REF. 1).
FT	CONFLICT	352	352		V -> E (IN REF. 1).
FT	CONFLICT	383	383		A -> T (IN REF. 1).
FT	CONFLICT	407	407		V -> I (IN REF. 1).
FT	CONFLICT	421	421		T -> S (IN REF. 1).
FT	CONFLICT	427	427		D -> N (IN REF. 1).
FT	CONFLICT	435	435		L -> V (IN REF. 1).
FT	CONFLICT	483	483		T -> I (IN REF. 1).
FT	CONFLICT	550	550		A -> V (IN REF. 1).
FT	CONFLICT	586	586		S -> L (IN REF. 1).
FT	CONFLICT	607	607		R -> H (IN REF. 1).
FT	CONFLICT	681	681		E9E8EDB29C79C450 CRC64;
SQ	SEQUENCE	695 AA;	78172 NM;		

[illegible]

CC 268 -----SLVYIAGLLAAAYOLVYGTGKRRF-----PPWLETLQCKKOLG--- 306
 CC Db 305 TQARGQVSLAEDGESSLAKEFDTWSEFGDYDLCSEVYDVI CSPEDDTENPCEDIDGHDI 364
 CC Qy 307 --LLSPFANVHA-----YGLCLPMRSEERYFLNMAVYQQVHANIESSME 351
 CC Db 365 LRVYIWIWISLITATGNIIVYVILTSQYKLTVE-----PFLMCLNLFAD----- 408
 CC Qy 352 EEVWRIEMYSFPGIMSGILSLLAVTSIPESVNALMRE-----PSFIQSTLG--- 399
 CC Db 409 -----LCIGIYLLILASVDIHKTQYHNAYIDMQGACDAGAFVTFASSELVYV 459
 CC Qy 400 YVALISTFH-----VLITGMKRAPEEERYRYPYTPNEF-LALVLP 439
 CC Db 460 LVTATLTERWHTTTHANOLQCKVKOLRHPASIMLVGMIFATFVLPFPFGISSYMKVSGICLP 519
 CC Qy 440 SYVLDLLOL 449
 CC Db 520 MDISP.LSQL 529
 CC
 CC RESULT 11
 CC SOTB_ECO57
 CC ID_SOTB_ECO57 STANDARD; PRT; 396 AA.
 CC AC P58529;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE Sugar efflux transporter.
 CC SOTB OR 22173 OR BCS2135.
 CC OS Escherichia coli O157:H7.
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 CC OX NCBI_TaxID=83334;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=O157:H7 / ATCC 700927;
 CC RX MEDLINE=21074935; PubMed=11206551;
 CC RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 CC Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 CC Postel G.G., Hackett J., Klink S., Boutin A., Siao Y., Miller L.,
 CC Raebrock E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K.,
 CC Apostolac J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 CC Welch R.A., Blattner F.R.;
 CC RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 CC RL Nature 409:529-533(2001).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=O157:H7 / RIMD 0509952;
 CC RX MEDLINE=21156231; PubMed=11258796;
 CC RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 CC Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 CC Rida T., Takami H., Honda T., Saekawa C., Ogasawara N., Yasunaga T.,
 CC Kubota S., Shiba T., Hattori M., Shinagawa G.
 CC RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 CC O157:H7 and genomic comparison with a laboratory strain K-12".
 CC RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: Involved in the efflux of sugars. The physiological role
 CC may be the reduction of the intracellular concentration of toxic
 CC sugars or sugar metabolites. Transports L-arabinose and to a
 CC lesser extent IPTG. Seems to contribute to the control of the
 CC arabinose regulon (by similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (probable).
 CC -1- SIMILARITY: Belongs to major facilitator superfamily. SotB
 CC (TC 2.A.1.2) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

```

CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL; AE005352; AAC56236.1; -
CC EMBL; AP002557; BAB35558.1; -
CC PIR; G90895; G90895.
CC PIR; H85721; H85721.
CC HAMAP; MF_00517; -; 1.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub transporter.
CC Pfam; PF00083; sugar tr; 1.
CC PROSITE; PS00850; MFS; 1.
CC Transport; Sugar transporter; Transmembrane; Inner membrane;
KM Complete proteome.
KM TRANSMEM 15 35 POTENTIAL.
KM TRANSMEM 50 70 POTENTIAL.
KM TRANSMEM 81 101 POTENTIAL.
KM TRANSMEM 103 123 POTENTIAL.
KM TRANSMEM 136 156 POTENTIAL.
KM TRANSMEM 170 190 POTENTIAL.
KM TRANSMEM 209 229 POTENTIAL.
KM TRANSMEM 246 266 POTENTIAL.
KM TRANSMEM 275 295 POTENTIAL.
KM TRANSMEM 299 319 POTENTIAL.
KM TRANSMEM 333 353 POTENTIAL.
KM TRANSMEM 364 384 POTENTIAL.
SQ SEQUENCE 396 AA; 42568 MW; CB6A209B0B5ED4F0 CRC64;

Query Match 4.4%; Score 104; DB 1; Length 396;
Best Local Similarity 18.8%; Pred. No. 1;
Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;

QY 31 VTGVTGSGDPAKSTIRLCRGYVVGSRNPKFSEFPFVVDVTHEDALTKNIIF 90
DB 31 VPTGGL--SDIAQSTHMOTAGVIMLTIVAVVVALMSLPFLMNSOVERKRLILCLEVF 88
QY 91 VAHREHTSLMDLPHLVGKILIDVSNRRINQDPESNAVLASLPDLSLVGFNVVS 150
DB 89 IASHVLSLTS--WSPFVLVISRI-----GVAFAAIF-----WSITA 123
QY 151 AMALQGP--KDSHQVYICSNIIQARQVIELAR-----OLNFIPIDLSL----- 195
DB 124 SLATMAAGKRAQSLIATGTAALAMVGLPGRIVGQYCGWMTFFALIGALITLCLC 183
QY 196 -----SSAREIENLPLRLFTLMRGVVAISLAT-----FFLVSEFVDVHPY 239
DB 184 LKILPLPSEHSGSLKSLPL-----LFRPALMSIYLLTVVVTVAHYTAYSY-----IEPF 235
QY 240 ARN---QSDPFKPIEIVNKTLPVATLTLSLVYLAGLAAAYQLYGTGRRFPWLE 296
DB 236 VQNTGFSANF-----ATALLLLGAGAILISVIFGKNGQYAS----- 274
QY 297 TWLQCRKQGLISFFPAAVVAVSYCLPFRSERVFLTMMAVQOVHANIEKSNKEEVR 356
DB 275 -----ATVSTAIALLVCLALLPAASE-----IHLGVLSIF-----WG 309
QY 357 IEMYSIFGMSLGLSL-----LAVTSLPSVSN-----ALNREFSFIQST 397
DB 310 IMMHTIGLMQVXVLAADADVDVMAALFSGIFNIGIGAGALVGNVLIHNS-----MSM 364
QY 398 LGVVALISTFHVLIYG-----WKRAPEEE 422
DB 365 IGVVG--TVPAFAALIMSIIFRRWPVLEEQ 394

RESULT 12
SOTB_SOTL
ID SOTB_SOTL STANDARD; PRT; 396 AA.
AC P31122; P76883; P77353;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sugar efflux transporter.
GN SOTB OR B1528.

```

```

OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=562;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayaishi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishino H., Nishio Y., Oshima T., Saito N.,
RA Sempel G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horinuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28,040,1 min region on the linkage map.";
RL DNA Res. 3:363-377 (1996).
RN (3)
RP SEQUENCE OF 217-396 FROM N.A.
RX MEDLINE=93186717; PubMed=8383113;
RA Cohen S.P., Haechler H., Levy S.B.;
RT "Genetic and functional analysis of the multiple antibiotic
RT resistance (mar) locus in Escherichia coli.";
RL J. Bacteriol. 175:1484-1492 (1993).
RN (4)
RP CHARACTERIZATION.
RC STRAIN=SB0;
RX MEDLINE=99194728; PubMed=10094697;
RA Bost S., Silva F., Belin D.;
RT "Transcriptional activation of ydeA, which encodes a member of the
RT major facilitator superfamily, interferes with arabinose accumulation
RT and induction of the Escherichia coli arabinose PBAD promoter.";
RL J. Bacteriol. 181:2185-2191 (1999).
RN (5)
RP CHARACTERIZATION.
RC STRAIN=US219;
RX MEDLINE=99369894; PubMed=10438792;
RA Carole S., Pichoff S., Bouche J.-P.;
RT "Escherichia coli gene ydeA encodes a major facilitator pump which
RT exports L-arabinose and isopropyl-beta-D-thiogalactopyranoside.";
RL J. Bacteriol. 181:5123-5125 (1999).
CC - FUNCTION: Involved in the efflux of sugars. The physiological role
CC may be the reduction of the intracellular concentration of toxic
CC sugars or sugar metabolites. Transports L-arabinose and to a
CC lesser extent IPTG. Seems to contribute to the control of the
CC arabinose regulon.
CC - SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC - SIMILARITY: Belongs to major facilitator superfamily. SotB
CC (TC 2.A.1.2) family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC
CC EMBL; AE000250; AAC74601.1; -
CC EMBL; D90795; BAA15210.1; -
CC EMBL; D90796; BAA15218.1; -
CC EMBL; D90797; BAA15230.1; -

```

DR EMBL: M96235; -: NOT_ANNOTATED_CDS.
 DR PIR: C64907; C64907.
 DR EcoGene; EGI1636; scb.
 DR HAMAP; MF_00517; -: 1.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR Pfam; PF00083; sugar tr; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR Transprot; Sugar transporter; Transmembrane; Inner membrane;
 KM Complete proteome.
 FT TRANSMEM 15 POTENTIAL.
 FT TRANSMEM 50 70 POTENTIAL.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 103 123 POTENTIAL.
 FT TRANSMEM 136 156 POTENTIAL.
 FT TRANSMEM 170 190 POTENTIAL.
 FT TRANSMEM 209 229 POTENTIAL.
 FT TRANSMEM 246 266 POTENTIAL.
 FT TRANSMEM 275 295 POTENTIAL.
 FT TRANSMEM 299 319 POTENTIAL.
 FT TRANSMEM 333 353 POTENTIAL.
 FT TRANSMEM 364 384 POTENTIAL.
 SQ SEQUENCE 396 AA; 42538 MW; C6A34CA4E6D4F0 CRC64;
 Query Match 4.4%; Score 104; DB 1; Length 396;
 Best Local Similarity 18.8%; Pred. No. 1;
 Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;
 QY 31 VTGVGIGSDPAKLTIRLCGVHVVIGSHNPKFSEFPFHVVDVTHEADLTQNIIF 90
 DB 31 VPTGL--SDAAGFHQTAQVIMLTIVAVVAMSLPFLMISQVRRRLICLFVVF 88
 QY 91 VAHREHTYSLMDRLHLVIGKILIDVSNMNEINQYPSNAEYLASLPDSLYKGVFVVS 150
 DB 89 IASHLSPLS--WSFVLVISRI-----GVAFAHIF-----WSITA 123
 QY 151 AMALQGP--KDSNQVYICNNICARQVIELAR-----QANFIPIDLSL----- 195
 DB 124 SLAIMRAGKRAQALSLIATGTALAWLGLPLRIYQVRCGMWTFPAIGLILILLC 183
 QY 136 -----SSARENLPRLFTLWFGPVVAISLAT-----FFLVSFVADVHPY 239
 DB 184 LKLPPLPSEHSGSLKSLPL-----LFRPALMSIYLLTVVVAHTAIVSY-----TEPF 235
 QY 240 ARN---QOSDPYKPIELIVNTLPVATLTLSTVYLAGLAAVQLVYGTYYRFPML 296
 DB 236 VQNIAGPSANF-----ATALLLLGAGIIGSVFGKLGNYAS----- 274
 QY 297 TWLOCRKQGLISFPFAMVHVAISICLPMSRSERYLFLNVAQQVHANINSENEEYVR 356
 DB 275 -----ALVSTAIALLVCLALLPRANSE-----IHGVLSIF-----WG 309
 QY 357 IEMYSFGIMSLGLSL-----LAVTSIPSVSN-----ALNREESFIQST 397
 DB 310 IAMMIIIGGVQVYALAPDADVMAALPSCGFNFICIGAGALVGNQVSLHMS-----XSM 364
 QY 398 LGYVALLISTFHALIYG-----WKRAREE 422
 DB 365 IGYVG-APDAFAALIMSIIFRRMPTLEEQ 394
 RESULT 13
 NPT1_MOUSE STANDARD; PRT; 465 AA.
 AC 061983;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Renal sodium-dependent phosphate transporter protein 1 (Sodium/phosphate cotransporter 1) (Na(+)/Pi cotransporter 1) (Renal sodium-phosphate transporter protein 1) (Renal Na(+)-dependent phosphate cotransporter 1).
 DE GN SLG17A1 OR NPT1.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=95335846; PubMed=7611445;
 RA Chong S.S., Kozak C.A., Liu L., Kristjansson K., Dunn S.T.,
 RA Bourdeau J.E., Hughes M.R.;
 RT "Cloning, genetic mapping, and expression analysis of a mouse renal
 sodium-dependent phosphate cotransporter."
 RL Am. J. Physiol. 268:F1038-F1045(1995).
 CC -1- FUNCTION: Important for the resorption of phosphate by the kidney.
 CC May be involved in actively transporting phosphate into cells via
 CC Na(+)-cotransport in the renal brush border membrane.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Kidney.
 CC -1- SIMILARITY: Belongs to the sodium/phosphate cotransporter family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X77241; CAA54459.1; -
 DR PIR: S69915; S69915.
 DR MGD; MGI:103209; Slc17a1.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR004745; Pi cotranspt.
 DR TIGRPFAM; TIGR00894; 2A0114euk.1.
 DR PROSITE; PS00850; MFS; 1.
 KW Transport; Symport; Sodium transporter; Transmembrane; Glycoprotein.
 FT TRANSMEM 79 99 POTENTIAL.
 FT TRANSMEM 117 137 POTENTIAL.
 FT TRANSMEM 176 196 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 260 280 POTENTIAL.
 FT TRANSMEM 304 324 POTENTIAL.
 FT TRANSMEM 337 356 POTENTIAL.
 FT TRANSMEM 363 383 POTENTIAL.
 FT TRANSMEM 399 419 POTENTIAL.
 FT TRANSMEM 429 449 POTENTIAL.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 47 47 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 56 56 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 465 AA; 51589 MW; C67BE25A2C291EEF CRC64;
 Query Match 4.4%; Score 103; DB 1; Length 465;
 Best Local Similarity 18.7%; Pred. No. 1.5; Indels 124; Gaps 16;
 Matches 76; Conservative 64; Mismatches 142;
 QY 110 GKILLIVSNMNIINQVPSNAEYLASLPDSLYKGVNVSAMALQGPDA---SRQY 166
 DB 79 GLTISVFFGQVAVAP---VGLSGIYPMKRIIGSLFSLMSMLLIPPAQVGAALVY 135
 QY 167 ICSNNTQARQVTELRQINFT---PIDLSGARGREINLPALFTIMRPVVY---- 218
 DB 136 VCRVLQGIAGVSTQGHIEIWMKMAPPLERGRILTS-----NTLSGFVN--GPFIVLVS 187
 QY 219 -----AISLATPFLVSFRVDVHPARQOSDF----- 247
 DB 188 GPICDLGLMPMVFYITGICVCLISWFLFFDDPRD--HPMSSSEKQYIISLMQAS 245
 QY 248 ---YKPIELIVNTLPDIVAIIIS----- 268
 DB 246 SGRQSLPIKMKSLPLMALIINSFAFIWSNSILVYTPFTISTYLHVNVRENGLSLIP 305
 QY 269 --LVYLAGLAAVQVYGTKYRRFPFMLETWQCRKQGLISFPFAMVHVAISICLPWR 326

DB 306 YLAVYICGILAGOMSPFLTRK-----IFSIVTRKLTTLGSCPVIFIMCLLYSYN 359

QY 327 RSEHYFLNMAVOVANIENSNNEEVRIRM---YISF--GIMSJL-----GLLSLAV 376

DB 360 FYSTVIFLTLA-----NSTLSFSCGOLNADLARIYGFILKAYALGIMGGLISSTLA 415

QY 377 TSPSVSNALNNEEFSFIQSTGLYVALISTFHVLT-----IYQWR 417

DB 416 GLINODPEYAMKISFLMAGINVTCLV---FYFLFANGELIDVAK 458

RESULT 14

NUCC_NEPOL STANDARD; PRT; 391 AA.

AC Q2RTK6;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE NAD(P)H-guinnone oxidoreductase chain H, chloroplast (EC 1.6.5.-)

DE (NAD(P)H dehydrogenase, chain H) (NADH-plastoquinone oxidoreductase 49 kDa subunit).

GN NDH.

OS Nephroselmis olivacea.

OC Chloroplast.

OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;

OC Chlorodendraceae; Chlorodendraceae; Nephroselmis.

OX NCBI_Taxid=11312;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NIES-464;

RX MEDLINE=99386894; PubMed=10468594;

RA Turrel M., Otis C., Lemieux C.;

RT "The complete chloroplast DNA sequence of the green alga Nephroselmis olivacea: insights into the architecture of ancestral chloroplast genomes."

RT Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).

CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED TO BE PLASTOQUINONE. COMPONENT OF THE IRON-SULFUR (IF) FRAGMENT OF THE ENZYME.

CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) + plastoquinol.

CC -1- SIMILARITY: Belongs to the complex I 49 kDa subunit family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF137379; AAD54891.1; -

DR InterPro; IPR001135; Oxidored 49kDa.

DR Pfam; PF00346; complex1_49kDa_1.

DR PROSITE; PS00535; Complex1_49K; 1.

KM Oxidoreductase; NAD; NADP; Guinnone; Plastoquinone; Chloroplast.

SO SEQUENCE 391 AA; 44913 MM; 9A40AEC6895825E CRC64;

Query Match 4.4%; Score 102.5; DB 1; Length 391;

Best Local Similarity 17.9%; Pctd. No. 1.3;

Matches 80; Conservative 70; Mismatches 124; Indels 173; Gaps 20;

QY 85 KTIIVFAIHHRYHTSLMDRLHLV---GKILID-----VSNMNMICYPE 127

DB 5 KIDPMIVSM-GPHHPMGVAKRLVITLDGENTLDECPVGVYHRMGEMIAENRTIVQYL 63

QY 128 --SNAEYLSLPFDSILVKGFNVSAMALQGPKASROVYICSNNICARQOVLARQL 185

DB 64 YVTRKMYLATMFTFETATV-----NABERLANTEVPRRA 96

QY 186 NRPIDIGLSARSARENLEPLRLFTMRGPVVVAISLATFFLVFVADVHPARMOQS 245

DB 97 SYLRVIMELSR-----IASHLWGPMDLGAQTEFFYLRRERMY----- 140

QY 246 DRYKPIEIVKTLPIVAITLTLVYLAAGLAAVGYGTRKRRPPMLETWLO-CRQ 304

DB 141 DLF-----AATGRMKNHYFRVGVYAA--DVPYG-----WIDKLDCEYR 180

QY 305 IGLSFFPAMV-----HVAYSICLPM-----RSEERYFL 334

DB 181 LPKVEYEALITRNPIFLKRVKGVITISPOQAINWGLSGEMLRASGVWDLRKVDYR-- 237

QY 335 NNAVQOVANIENSNNEEVRIM-----EYISFGISLGLSLAVTSIP----- 380

DB 238 -ECYEDFHWVSSESTGCLARVLRKERTSTIKYQ-----QAKSLPGSTENLEA 230

QY 381 --SVSNALNNEEFSFIQSTGLYVALISTFHVLT-----IYQWR 415

DB 291 RQLSGRTSPNNEFDPYQFLGKASPTKPMRQEHYVVEAKBELGVFLIGDHPFMRW 350

QY 416 KRAFEERYRFTYPPNLYALVPSIV 442

DB 351 K-----IRPRGFINQILPNTV 367

RESULT 15

Y304_BRUME STANDARD; PRT; 220 AA.

AC Q8YD73;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein BME10304.

DE BME10304.

GN BME10304.

OS Brucella melitensis.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Brucellaceae; Brucella.

OX NCBI_Taxid=29459;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=16M / ATCC 23456 / Biotype 1;

RX MEDLINE=20020109; PubMed=11756688;

RA Delvecchio V.G., Kapetral V., Reckar R.J., Patra G., Wajer C., Los T., Ivanova N., Anderson I., Bhattacharya A., Lydids A., Reznik G., Jabloncki L., Larsen N., D'Souza M., Bernal A., Goldsman E., Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J., Haselkorn R., Kyrides N., Overbeek R.;

RT "The genome sequence of the facultative intracellular pathogen Brucella melitensis."

RT Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: Belongs to the UPF0191 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AE009669; AAL53546.1; -

DR PIR; AG3547; AG3547.

DR HAMAP; MF_01207; -; 1.

DR InterPro; IPR007916; UPF0191.

DR Pfam; PF05252; UPF0191; 1.

KM Hypothetical protein; Transmembrane; Complete proteome.

SO SEQUENCE 220 AA; 24815 MM; 182C0244743B17FA CRC64;

Fri Apr 16 16:26:36 2004

us-09-455-486-6.rsp

Page 13

Query Match	Score	ID	DB	Length
4.3%	102	1	220	
3.8%	102	1	220	
3.8%	102	1	220	

Best Local Similarity 23.8%; Pred. No. 0.72;
Matches 43; Conservative 25; Mismatches 51; Indels 62; Gaps 7.

267 LSLVYLGLIAAAYQLYYGT-----KYRRFPWLETW----- 298

Db 20 LMLLTAGFVPAVWTFYLGATGQLGADPAVKTFEHLGLWALRFLITLLVTPMRDLTGIT 79

299 -JQCRKOLGILSFFAMVHVAYSCLPNERSERYLFIMAYQVHANIEINSWNEEBVRI 357

Db 80 LIRYRRALGLAFYYALMHT-----TYMVLDOGLNSAITDIYR- 120

QY 358 EWISFGIMSLGLSLAVTSPISVSNAIWN-----RPSFIQSTLGLYALLISTFHWL 411

Db 121 RPFITGMISLALVPLALTS-----NNWSIRKLGRWSSLHK-LVYIAIAGSAVHFL 172

QY 412 I 412

Db 173 M 173

Search completed: April 16, 2004, 12:39:41
Job time : 20 secs

Search completed: April 16, 2004, 12:39:41
Job time : 20 secs

AC2560
 hypothetical protein alr8074 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120ga
 C/Species: Nostoc sp. PCC 7120
 A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C/Accession: AC2560
 R/Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchih
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
 A/Reference number: AB1807; MUID:21595285; PMID:11759840
 A/Accession: AC2560
 A/Molecule type: DNA
 A/Status: preliminary
 A/Residues: 1-211 <KIR>
 A/Cross-references: GB:AP003603; PIDN:BA877404.1; PID:g17134847; GSPDB:GN00182
 A/Experimental source: strain PCC 7120
 C/Genetics:
 A/Gene: alr8074
 A/Genome: plasmid

Query Match 6.8%; Score 159.5; DB 2; Length 211;
 Best Local Similarity 24.9%; Pred. No. 2.4e-05;
 Matches 51; Conservative 43; Mismatches 70; Indels 41; Gaps 8;

QY 33 VGVIGSGDFAKSLTIRLCGYHVVIGSRNP-----KFASEFPFHVVDVTHEDALTKTN 87
 DB 3 IAFIGIGVGSALASQLSLDHTVTIAARNSNSDSVKALAKYRP-LQVSSPQELIAQAE 61
 QY 88 IIFVIA-----IHEHTSLMDLRLHLVGLKILIDVSNMNRIN-----QYESNAEYLALSLFP 138
 DB 62 VIFLPTPLFPAQALAEVDDL-----GKLVDCNTPVGVNLTHGKSGSSSELVQSVVP 117
 QY 139 DLIYKGFNVVSAWALQIGPKASROY-----ICSNIOARQOVIETLARQINF 187
 DB 118 HAKVYKAFIT-----YGFENFENHTYPGYGNLKPAMLIAGNDVPAQVAVSLCQQLGWM 170
 QY 188 IPIIDGSSAREIENLPLRLFTIWM 212
 DB 171 EAVDVGNSLSLHLEHTL-----LM 191

RESULT 3

TS0571
 probable oxidoreductase [imported] - Streptomyces coelicolor
 C/Species: Streptomyces coelicolor
 C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 28-Jul-2000
 C/Accession: TS0571
 R/Federbach, M.; Kleser, H.M.; Denapate, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
 Mol. Microbiol. 21, 77-96, 1996
 A/Title: A set of ordered cosmid and a detailed genetic and physical map for the 8 Mb S
 A/Reference number: Z20556; MUID:97000351; PMID:8843436
 A/Accession: TS0571
 A/Status: preliminary; translated from GB/EMBL/DDJ
 A/Molecule type: DNA
 A/Residues: 1-239 <RED>
 A/Cross-references: EMBL:AL133220; PIDN:CA861708.1
 A/Experimental source: strain A3(2)
 C/Genetics:
 A/Note: SC075A.08c
 C/Superfamily: conserved hypothetical protein MJ1501

Query Match 6.7%; Score 156.5; DB 2; Length 239;
 Best Local Similarity 27.9%; Pred. No. 5e-05;
 Matches 61; Conservative 41; Mismatches 72; Indels 45; Gaps 10;

QY 25 IKDARKTYGVI-GSGDFAKSLTIRLCGYHVVIGSR--NPKFASEFPFHVVDVTHHE 80
 DB 23 LPDVGSLVGVLTGSGTQGGKGLARLAKAGQKIVGSRABRAAAAEIGHGVGADNA 82
 QY 81 DALTKNIIFVAIHR- HTYSLMDLRLHLVGLKILIDVSNMNRINQ-----YFE--SNE 131
 DB 83 ETARSDVIVAVPMDGHGKTLISRLAELSGKLVVDCVNPFGDKGAVALKPEBSAAB 142

QY 132 YIASLFPDSLIVGFNVVSAWALQ-----LGPKDASROYVIGSNNOARQOV 178
 DB 143 QAAALLPDSRVAAAFVHLSAVLLOPEIDEIDTVWVLGERRADVEI-----VOA----- 192
 QY 179 IELAQQLNFI-----IDIGSSAREIENLPLRLFTIWM 213
 DB 193 --LAGR---IPGMRGVAFGRLLNHAQVBSLVANLISVNR 226

RESULT 4

conserved hypothetical protein AF1209 - Archaeoglobus fulgidus
 C/Species: Archaeoglobus fulgidus
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
 C/Accession: H69400
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A/Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A/Reference number: A69250; MUID:98049343; PMID:9389475
 A/Accession: H69400
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-213 <KLE>
 A/Cross-references: GB:AE001021; GB:AE000782; NID:G2689344; PIDN:AA89038.1; PID:g264937
 C/Superfamily: conserved hypothetical protein MJ1501

Query Match 6.3%; Score 147; DB 2; Length 213;
 Best Local Similarity 27.5%; Pred. No. 0.0024;
 Matches 56; Conservative 44; Mismatches 68; Indels 36; Gaps 11;

QY 37 GSGDFAKSLTIRLCGYHVVIGSRN-----KFASEFPFHVVDVTHEDALTKTN 88
 DB 8 GTNIGEGALRLRGKUGYELIVGSRKLEKAKLASDYLKVGDAISIGKENDDAETCD- 66
 QY 89 IFAIHEHTSLMD-----LRHLVGLKILID-----VSNMNRINQYPE--SNAE 131
 DB 67 --VAV-----FTIPWEFAFDTAEMLKQLAGKVVISPLVPMKRYGDNF-VYVPEEGSABE 119
 QY 132 YIASLFPDSLIVGFNVVSAWALQIGPKASROYVIGSNNOARQOVIETLARQINF-P1 190
 DB 120 KLASVLEESSVVAAYHSIPARRPANTGEEFEMVPICGDS-GAKYVVDLTKISGLRL 178
 QY 191 DLGSSAREIENL-PLRLFTIWM 213
 DB 179 DAGLSNAHLVESITPLINVMKR 202

RESULT 5

D69361
 conserved hypothetical protein AF0892 - Archaeoglobus fulgidus
 C/Species: Archaeoglobus fulgidus
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
 C/Accession: D69361
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A/Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A/Reference number: A69250; MUID:98049343; PMID:9389475
 A/Accession: D69361
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-212 <KLE>
 A/Cross-references: GB:AE001042; GB:AE000782; NID:G2689365; PIDN:AA890348.1; PID:g26497
 C/Superfamily: conserved hypothetical protein MJ1501

Query Match 6.0%; Score 142; DB 2; Length 212;
 Best Local Similarity 25.4%; Pred. No. 0.0006;
 Matches 49; Conservative 44; Mismatches 84; Indels 16; Gaps 6;

QY 37 GSGDPAKSLTIRLRGCVHVVIGSRNPKF---FASEFPFHVVDV-----HHEDALTKNI 88
 DB 8 GTGDLGGLALRLATLGHEIVSGRRREKAEAKAEARRIAGDASITGMKNEDAAACDI 67
 QY 89 IFVAIHREH-YTSLMDLRHLVGIKII---LIVSNMKRINQY--PESNAEYLASFPPDSL 141
 DB 68 AVLTIPWEHAIIDTARDLKNILREKIVSPVPSRGAKGFTYSERSAAEIVAEVSESEK 127
 QY 142 IVKGFNVASAMALQGPKASROYVICSNNIQARQVIELARQINFI-PIDLSISAPAE 200
 DB 128 VVSALHTIPARFANLDEKFDVPCVGGDDSEKVMSLSELGRLPDAFPLNSRL 187
 QY 201 IENLPRLFTLMR 213
 DB 188 VESLTPLINIMR 200

RESULT 6

T10120
 F420-dependent NADP reductase (EC 1.6.8.-) - Methanobacterium thermoautotrophicum (strain
 C/Species: Methanobacterium thermoautotrophicum
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 C/Accession: T10120
 R/Berk, H.; Thauer, R.K.
 FEBS Lett. 438, 124-126, 1998
 A/Title: F420H2:NADP oxidoreductase from Methanobacterium thermoautotrophicum: identific
 A/Reference number: Z16959; MUID:99037734; PMID:9821972
 A/Accession: T10120
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-224 <BBP>
 A/Cross-references: EMBL:Y17210
 A/Experimental source: Strain Marburg
 C/Superfamily: conserved hypothetical protein M01501
 C/Keywords: oxidoreductase

Query Match 6.0%; Score 142; DB 2; Length 224;
 Best Local Similarity 29.2%; Pred. No. 0.00064;
 Matches 59; Conservative 29; Mismatches 84; Indels 30; Gaps 8;

QY 37 GSGDPAKSLTIRLRGCVHVVIGSRNPKFSEFPFHVVDVTHHEDALTK--TN----- 87
 DB 8 GTGDLGGLALRLATLGHEIVIGSRDAKAVSAQKVLIERDRLKVGATNAEAAEEA 67
 QY 88 ---IIFVAIHREHYSMDLRHLVGIKIIIV-----SNMKRINQY--PESNAEYL 134
 DB 68 EVALTIPWEHAIIDTARDLKNILREKIVSPVPSRGAKGFTYSERSAAEIVAEVSESEK 126
 QY 135 SLFPD-SLIVKGFNVASAMALQ-LGPKDASROYVICSNNIQARQVIELARQINFI-P 189
 DB 127 RFLPDQGTVAALFANNISASALLDITGPVDC--DCLIASDRDALDLASELAKIDGVRA 184
 QY 190 IDLSISASAREIENLPRLFTL 211
 DB 185 IDCGLENARVIEKITPLINL 206

RESULT 7

D64487
 Hypothetical protein M01501 - Methanococcus jannaschii
 C/Species: Methanococcus jannaschii
 C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C/Accession: D64487
 R/Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinrock, K.G.; Merrick, J.M.; Glodek, A.;
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Frazer, C.M.; Smith, H.O.; Woese, C
 A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A/Reference number: A64300; MUID:96337999; PMID:8688087

A/Accession: D64487
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-223 <BBU>
 A/Cross-references: GB:U67591; GB:L77117; NID:G2826422; PIDN:AA89514.1; PID:G1500389; T
 C/Genetics:
 A/Map position: REV1473617-1472946
 C/Superfamily: conserved hypothetical protein M01501

Query Match 5.7%; Score 133; DB 2; Length 223;
 Best Local Similarity 21.9%; Pred. No. 0.0033;
 Matches 46; Conservative 52; Mismatches 78; Indels 34; Gaps 7;

QY 37 GSGDPAKSLTIRLRGCVHVVIGSRNPKFSEFPFHVVDV-----THHEDALT 84
 DB 8 GTGDLGGLALRLAK--NKKIIGSRKKEAEAKAEKIKORGIEADITGLEKDAK 66
 QY 85 KNIIFVAIHREH-YTSLMDLRHLVGIKIIIV-----SNMKRINQY--PESNAEYL 133
 DB 67 EGDVILSLPYEYTLSTIKOLKEELKGIIVSIGVPLATAGDKPTRLFPDGSVAEMV 126
 QY 134 ASLPPDSLIVKGFNVASAMALQGPKASROYVICSNNIQARQVIELARQINFI-PIDL 192
 DB 127 QNVLESKYVAFQVCHAVLEDDNPVDCDILVGNDEAKKVVYDLANGIDGVRAIDC 186
 QY 193 GSLSAREIENLPRLFTLMRGPVVAISL 222
 DB 187 GNLEKSRITAEIT-----PLDIGNI 207

RESULT 8

F86826
 Hypothetical protein ygfE (imported) - Lactococcus lactis subsp. lactis (strain IL1403)
 C/Species: Lactococcus lactis subsp. lactis
 C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C/Accession: F86826
 R/Bolotin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrh,
 Genome Res. 11, 731-753, 2001
 A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
 A/Reference number: A86625; MUID:21235186; PMID:11337471
 A/Accession: F86826
 A/Status: preliminary
 A/Experimental source: strain IL1403
 A/Residues: 1-191 <STO>
 A/Cross-references: GB:AE005176; PID:G12724621; PIDN:AAK05712.1; GSPDB:GN00146
 A/Experimental source: strain IL1403
 C/Genetics:
 A/Gene: ygfE

Query Match 5.1%; Score 121; DB 2; Length 191;
 Best Local Similarity 22.2%; Pred. No. 0.024;
 Matches 47; Conservative 33; Mismatches 74; Indels 58; Gaps 7;

QY 32 TVGVIGSGPAKSLTIRLRGCVHVVIGSRNPKFSEFPFHVVDVTHHEDALTKTN----- 87
 DB 3 TISIFGKXGKXKAI-----GNFSSVAKVYITLNSKTELGRI 42
 QY 88 -----IIFVAIHREHYSMDLRHLVGIKIIIVSNMKRINQY-----SNAEYL 133
 DB 43 VVLAVPYVALAGIIGVSTD-----LQKIIITDITNPVDFTTSPSLVPSPTSAALII 95
 QY 134 ASLPPDSLIVKGFNVASAMALQGPKASRQ-----YVICSNNIQARQVIELARQINFI-P 189
 DB 96 AKOLPMSIMYKAFN--TFESDTATKVAANEHQTIVLLASDSQAEKTIKALENGSLI 153
 QY 190 IDLSISASAREIENLPRLFTL-----WRG 214
 DB 154 LDAGSLRARELEAIGLQITTLASEKISIDG 185

RESULT 9

A69131

conserved hypothetical protein MTH248 - Methanobacterium thermoautotrophicum (strain Del
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C:Accession: A69131
R:Smith, D.R.; Doucette-Stamm, L.A.; DeJonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadotora, R.; Viscare, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: A69131
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-232 <MTH>
A:Cross-references: GB:AE000811; GB:AE000666; NID:92621287; PIDN:AA84754.1; PID:9262128
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH248
A:Start codon: TTG
C:Superfamily: conserved hypothetical protein MJ1501

Query Match 5.1%; Score 120.5; DB 2; Length 232;
Best Local Similarity 23.8%; Pred. No. 0.034;
Matches 49; Conservative 42; Mismatches 90; Indels 25; Gaps 6;

30 KATVGVIGSDPAKSLTIRIRCGYHVIGSRPKFSEFPFHVVDTHHE----- 80
DB KIN-IGTGDDQGLALRFVAVGEVITIGSRRAEKASVAGVLEIARDISVEGATN 66
81 -DALTKNITFVAIH-REHYTSLMDRLHVLGKILIDV-----NNRINQYPSRN 129
DB PDAAASADVAVLTPLOQWVTLASRDQVRDVLIDAVPDISCIGGAVYIDLMGSG 126
130 AEYASIF---PDLIKYGNVVSAMALQGPXASROYVIGSNNOARQVIELAROLN 186
DB AAEEAARFIREQGRVAAAFNNISASALLESSEPVDCCLVNSDHDALVAAELAEKID 186
127
187 FI-PIDIGSSASAEIENPLRFTL 211
DB GVRALIECGLENNARILEKTPPLINTL 212

RESULT 10
T00121
hypothetical protein 8 - Leptospira interrogans

C:Species: Leptospira interrogans
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00121
R:Itakashi, Y.; Akase, K.; Hirano, H.; Fukunaga, M.
Gene 215, 37-45, 1998

A:Title: Physical and genetic maps of the Leptospira interrogans serovar Icterohaemorrhagiae
A:Reference number: Z14115; MUID:98327217; PMID:9666070
A:Accession: T00121
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-216 <TK>
A:Cross-references: EMBL:AB010203; NID:92780763; PIDN:BA24368.1; PID:92780765
A:Experimental source: strain Ictero No.1; substrain Icterohaemorrhagiae

Query Match 5.1%; Score 119.5; DB 2; Length 216;
Best Local Similarity 23.7%; Pred. No. 0.037;
Matches 47; Conservative 37; Mismatches 91; Indels 23; Gaps 6;

33 VGVIGSDPAKSLTIRIRCGYHVIGSRPKFSEFPFHV---VDVTHEALTKNTI 89
DB IGIIGSGIVGGLTANGLKGAAYKIGTRDFGLKDLAKAGASIGSFSSANFGEII 65
90 FVAIHREHYTSLMDRLH---LVGKILIDVSNRR-----INQPSNAAVYASL 136
DB VLCKSGVAVSEVLTLSGIDSLNGKTIIDTNPISERIPQNGVNFPTSNESIMKELQK 125
137 FPDLSIKGFNVVSAMALQGP--KDSROYVIGSNNOARQVIELAQLNFIPIDLOS 194

DB 126 AKRANVKKFSSVGS-GLVNPDQLKGEKPSWIFIGNDDSSAKQKELIDFGMDTEWGMK 184
CY 195 LSGAREIENLPRLTTLW 212
DB 185 VEAARIEPLCT---LW 198

RESULT 11
G82642
conserved hypothetical protein XF1737 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: G82642
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below

A:Accession: G82642
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-242 <SIM>
A:Cross-references: GB:AE003997; GB:AE003849; NID:99106805; PIDN:AAE84546.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Brienes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carver, F.
as-Melo, E.; Docena, C.; El-Dorry, H.; Facincenti, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Froh
J.D.; Jungueira, M.L.; Kemper, E.L.; Kitzajima, J.P.; Krieger, J.E.; Kurame, E.B.; Laiz
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E.
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasat
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tashiro, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovsky-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1737

Query Match 5.0%; Score 116.5; DB 2; Length 242;
Best Local Similarity 24.0%; Pred. No. 0.075;
Matches 46; Conservative 40; Mismatches 85; Indels 21; Gaps 7;

CY 28 AKRANVKKFSSVGS-GLVNPDQLKGEKPSWIFIGNDDSSAKQKELIDFGMDTEWGMK 83
DB 30 AADPRIGVIGASLGGTGRVWVKGHEWVPSRVPDLAMAELEBR-ASVQPIAAI 88
CY 84 TKNITFVAIHREHYTSLM-DRLHVLGKILIDVSNRRINQ---YPSN---AEYIAS 135
DB 89 EFGTVLLAVPEALPQVRDLSAYRGKIVLDSNPGASADVYREARELGAQVTVK 148
CY 136 LFPDLSIKGFNVVSAMALQGPXASROYVIGSNNOARQVIELAROLNFIPI 190
DB 149 YMGARLVAFSAVDITVY---TSASRGGRIGMPLASDDAEAMKVAEGVYRAGCDPV 205

CY 191 DLGSSSAREIE 202
DB 206 IVGNLAAASFO 217

RESULT 12

AB3182
conserved hypothetical protein Atus183 [imported] - Agrobacterium tumefaciens (strain C
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AB3182
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
rase, G.; Gallet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCell
Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
 A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB25777; MUID:21608550; PMID:11743193
 A:Accession: AB3182
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-198 <KUR>
 A:Cross-references: GB:AE008687; PIDN:AA45872.1; PID:g17743615; GSPDB:GN00188
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: At5g183
 A:Genome: plasmid

Query Match 4.9%; Score 116; DB 2; Length 198;
 Best Local Similarity 22.7%; Pred. No. 0.063;
 Matches 50; Conservative 41; Mismatches 81; Indels 48; Gaps 10;

QY 31 VTVGVIGSGDFAKSLTIRLCGYHVI-GSRNPKFASEFPVH--VDVTHHEDALTKT 86
 DB 1 MTVGIIIGAGNIGAFATALGKAGIEAVIANSRGPESLTAIVSKGSTRAGSVPEAAQA 60
 QY 87 NIIVVAIHRREHYTSLMD-----LEHLVG-KLIIIVSNMNRINOY-----PESNAEYLA 134
 DB 61 -IVVAVP-----WSKIPGALAGLNFGRIVIDANNISIEAPLYRPADIGRTSDIFT 112
 QY 135 SLFPDSLVKGFNVVSAVALQIGPKDAS-----ROYVICSNNIQAARQVITELAROL 185
 DB 113 ALVPGARVAVKAFN-----HLTEKQLSGDPHSEGRVLFYSGDDWRKAEVGAIIIDRI 165
 QY 186 NFIPIDIGSLSSAREIENLPRLFTLWGRPVVAISLATP 225
 DB 166 GFGEIDLGLEFVGSQLQFP-----GGFPLALNLVKF 197

RESULT 13

JN0898
 A:Alternate names: follicle-stimulating hormone receptor (FSHR)
 C/Species: Macaca fascicularis (craab-eating macaque)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: JN0898; S36452
 R:Gromoll, J.; Dankbar, B.; Sharma, R.S.; Nieschlag, E.
 Biochem. Biophys. Res. Commun. 196, 1066-1072, 1993
 A>Title: Molecular cloning of the testicular follicle stimulating hormone receptor of th
 A:Reference number: JN0898; MUID:94071854; PMID:7504463
 A:Accession: JN0898
 A:Molecule type: mRNA
 A:Residues: 1-695 <GRO>
 A:Cross-references: EMBL:X74454; NID:9396801; PIDN:CAA52463.1; PID:9396802
 A>Note: the authors translated the codon AGT for residue 488 as Arg
 C/Function:
 A:Description: receptor that mediates the biochemical effects of follicleotropin
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotein;
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:16-95/Domain: signal sequence #status predicted <SIG>
 F:16-95/Product: follicleotropin receptor #status predicted <PFR>
 F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:146-168/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F:367-387/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F:399-421/Domain: transmembrane #status predicted <TM1>
 F:444-465/Domain: transmembrane #status predicted <TM2>
 F:486-508/Domain: transmembrane #status predicted <TM3>
 F:529-550/Domain: transmembrane #status predicted <TM4>
 F:574-597/Domain: transmembrane #status predicted <TM5>
 F:609-630/Domain: transmembrane #status predicted <TM6>
 F:191,199,293,318/Binding site: carboxylate (Asp) (covalent) #status predicted
 F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 4.7%; Score 111.5; DB 1; Length 695;
 Best Local Similarity 17.0%; Pred. No. 0.071;
 Matches 84; Conservative 75; Mismatches 151; Indels 183; Gaps 16;

QY 79 HEDALTKNIIIVAIHREHYTSLMDRLHLVG-----KLIIIVSNMNR 121
 DB 98 HEIRIEKANNL-LVINPEAFQNLPLNRYLLISNTGIKHLPDVKHISFOKVLIDIQNIN 156
 QY 122 INQVESAAYLASLFPDSLVKGFNVVSAVALQIGPKDASROYVICSNNIQAARQVIEL 181
 DB 157 IH-----TIERNSFVGLSPESVITML-----NKNGIQEIHNCA----- 189
 QY 182 AROLNFIPIDIGSLSSAREIENLPRLFTLWGRPVVAISLATPFFYSGVDVHYHYPAR 241
 DB 190 ---FNGTQDELINLSDNNNLDELNDVDFHAGSGEVIIDISRTIRHSIPSGLENLKCLRA 246
 QY 242 NQGSDFYKIPPIEIVNKLPIVAITLLSVY----- 271
 DB 247 RSTYVTKLP-----SLEKVALMEASLTYPSCCAFANMRQISELHPICNKSILROEV 301
 QY 272 -----LAGLLAAAYQLYYGYKRRF-----PWLLETWLQCRKQLG 306
 DB 302 DYMOTRGQSSLAEDNESSSRGDMTYAEFDYDLCEVVDVTCSPKPDAFNFCEDILG 361
 QY 307 ---LSFFPAMHVA-----YSCLPFRSRRLPLMNAQQVHANIEINS 348
 DB 362 YNLRVLIWFISILATIGNIIVLVTTSQKLTVP-----RFLMCLNLFAD----- 408
 QY 349 WNEEWEWRIEWISFGIMSLGLSLAVTSIPSVSNALNMBE-----FSFIQSTLG 399
 DB 409 -----LCGIYLLLIASVDYHTKQYHNVAIDWOTGAGCDAGFFTFASGLS 456
 QY 400 ---YVALLISTFH-----VLIGKRAFPBEHYVFPFV-LAL 436
 DB 457 VYTLTALITLRRHTTTHAMQDLCKVHVAASVWVGWIFAPAAALFPIFGISSTKYSI 516
 QY 437 VPSIVILDLQL 449
 DB 517 CLPMDISPLSQL 529

RESULT 14

S59681
 A:probable membrane protein YP012w - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein LP45w, hypothetical protein YP8132.01
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 15-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002
 C/Accession: S59681; S52519
 R:Hall, J.; Ahmed, A.; Busssey, H.; Fortin, N.; Friese, J.D.; Storms, R.K.; Vo, D.H.; Wa
 submitted to the EMBL Data Library, August 1995
 A:Description: The sequence of *Saccharomyces cerevisiae* chromosome XVI left arm.
 A:Reference number: S59677
 A:Accession: S59681
 A:Molecule type: DNA
 A:Residues: 1-1228 <HAL>
 A:Cross-references: EMBL:U3335; NID:9365076; PID:9365081; MIPS:YP012w
 R:Badcock, K.; Chutcher, C.
 submitted to the EMBL Data Library, February 1995
 A:Reference number: S52519
 A:Accession: S52519
 A:Molecule type: DNA
 A:Residues: 220-1228 <BAD>
 A:Cross-references: EMBL:248483; NID:96683777; PID:96683778; MIPS:YP012w
 A:Experimental source: strain AB972
 C:Genetics:
 A:Gene: SGD:RRP12
 A:Cross-references: SGD:S0005933
 A:Map position: 16L
 C:Keywords: transmembrane protein
 F:135-141/Domain: transmembrane #status predicted <TM1>
 F:478-494/Domain: transmembrane #status predicted <TM2>
 F:720-736/Domain: transmembrane #status predicted <TM3>

Query Match 4.7%; Score 111.5; DB 2; Length 1228;

Best Local Similarity 19.9%; Pred. No. 1.5;

Matches 111; Conservative 71; Mismatches 152; Indels 225; Gaps 28;

QY 11 KLSLSETCPLNGINGIKDARK-----VTWVGISGDFAKSLTIRLI 50

DB 292 KTMLETITSSLAENKIRVDTITFALKPSVNDLLTKSWIAVYIKMSYATCQPLKAL 351

QY 51 R-----CGYHVGISNPKF--ASEFPVAVDVTTHEDL----- 83

DB 352 RKIPGVHIMCTY--LASETPEVYQAASQCLISILSESVDDLLVTPSVDEKVRND 408

QY 84 -----TKTNIIFVAIHREHYTSLMDLHLVKGILIPVSNMNINGPSNAEYLAFL 137

DB 409 EIIISQIAKTFTDLISIRSHCS-----REIL--KILVAAPKFRY--RNPHTLSL- 456

QY 138 PDSLIVKGFNVSAVALQGPXASRQVYICSNNIQARQVIELARQINFI----- 188

DB 457 -----KIYDTWRVN-----EEOFMDLRNEIELVIGASISAMG 488

QY 189 -----PIDGSSAREIENLPLRLTLRGPV-----VAISLAFPLYSFVRDVI 236

DB 489 PEMTIAEAPLNDNPSER-----PGR--AMLPILRDYTKNATLAF-----QNEL 533

QY 237 HPEANQOSDFYKIPLEIYN-----KTLF-----TVAITLTL 269

DB 534 AVYIKSPQSKFQVPEESQLKVFQTIVDQIWSTLPRCELPMDLRESFTDEFASLSTL 593

QY 270 VTL-----AGLLAAAYQLYGTR-----YRFP-----PWLEMTLQCRK 303

DB 594 LYSEVELRTTICHLKVAESNVSYAESSSHNVLLOFPISBAQNIETYST-----K 648

QY 304 QGLSFPFAMVVAVSLCPMERS-----ERYL-----FLNMAVOOVANIENSUNE 351

DB 649 STNLLAVLFNV---YTQTPNARSYLETIDQYIKTSKEDLEKTFNNVCGELKSNVNE 704

QY 352 EEWRIEMTISFGINSGLSL--AVTSPVSNAIMRERSFIQSTLGYVALLISTFH 409

DB 705 EESGNVNERKKRPQLATLIDLICITLYLPVSYISALFSNFS-----LTVNSAD 754

QY 410 VLIYGWKRABEZYRYFYT 428

DB 755 ALIQ--KRA-----YRIIT 766

RESULT 15

B84410 hypothetical protein Vng2607c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C:Accession: B84410

R:NG, M.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

; Leitbauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li

A>Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; PMID:20504483; PMID:11016950

A:Accession: B84410

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-222 <STO>

A:Cross-references: GB:AE004437; NID:G10581987; PIDN:AA020646.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG2607C

C:Superfamily: conserved hypothetical protein MJ1501

Query Match 4.7%; Score 110.5; DB 2; Length 222;

Best Local Similarity 24.3%; Pred. No. 0.2;

Matches 49; Conservative 34; Mismatches 94; Indels 25; Gaps 6;

DB 8 GTGDIGAGLALRMATDSHDIVISGRDEKARETAAYEDTLADQVDRKLTGFANEMAA 67

QY 84 TKTNIIFVAIHREHYTSLM-----DLRHLVKGILIPVSNMNINGPSNAEYLAFL 133

DB 68 DRADVAVVAVPAYHVTVDWGAVALRDLADDTLVISPAVGIASGEHGHYNPSSAGSVTALV 127

QY 134 ASLPPDSL-IVKGFNVSAVALQGPXASRQVYICSNNIQARQVIELARQINFI-PID 191

DB 128 ADAAPDGVAVGAFINLAADRLADDTLVDGNDGARTRVABLADDTGLRALD 187

QY 192 LGSLSAREIENLPLRLTLR 213

DB 188 AGPVENMAEVESLTPLINLAR 209

Search completed: April 16, 2004, 12:41:12

Job time : 23 secs

Fri Apr 16 16:26:35 2004

US-09-455-486-6.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:40:42 ; Search time 48 Seconds

(without alignments)
2607.409 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351

Sequence: 1 MESISMWGSFKSLSETCLPN.....ALVPSIVLDLQICRYDP 454

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: Published Applications_AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pdb:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pdb:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pdb:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pdb:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pdb:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pdb:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pdb:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pdb:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pdb:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pdb:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pdb:*
14: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pdb:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pdb:*
16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pdb:*
17: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pdb:*
18: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pdb:*
19: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2351	100.0	454	US-09-888-257A-10	Sequence 10, Appl
2	2351	100.0	454	US-10-455-822-3	Sequence 3, Appl
3	2351	100.0	454	US-10-455-822-7	Sequence 7, Appl
4	2351	100.0	454	US-10-455-822-9	Sequence 9, Appl
5	2351	100.0	454	US-10-455-822-19	Sequence 19, Appl
6	2351	100.0	454	US-10-455-822-21	Sequence 21, Appl
7	2351	100.0	454	US-10-455-822-23	Sequence 23, Appl
8	2351	100.0	454	US-10-455-822-25	Sequence 25, Appl
9	2351	100.0	454	US-10-455-822-31	Sequence 31, Appl
10	2351	100.0	454	US-10-455-822-33	Sequence 33, Appl
11	2351	100.0	454	US-10-455-822-35	Sequence 35, Appl
12	2351	100.0	454	US-10-455-822-37	Sequence 37, Appl
13	2351	100.0	454	US-10-455-822-39	Sequence 39, Appl
14	2351	100.0	454	US-10-455-822-78	Sequence 78, Appl
15	2351	100.0	454	US-10-455-822-121	Sequence 121, Appl

15	2351	100.0	454	US-10-455-822-160	Sequence 160, Appl
16	2351	100.0	454	US-10-455-822-161	Sequence 161, Appl
17	2351	100.0	454	US-10-455-822-162	Sequence 162, Appl
18	2351	100.0	454	US-10-455-822-163	Sequence 163, Appl
19	2351	100.0	454	US-10-455-822-164	Sequence 164, Appl
20	2351	100.0	454	US-10-455-822-165	Sequence 165, Appl
21	2351	100.0	454	US-10-455-822-166	Sequence 166, Appl
22	2351	100.0	454	US-10-455-822-167	Sequence 167, Appl
23	2351	100.0	454	US-10-455-822-168	Sequence 168, Appl
24	2351	100.0	454	US-10-455-822-169	Sequence 169, Appl
25	2351	100.0	454	US-10-455-822-170	Sequence 170, Appl
26	2351	100.0	454	US-10-455-822-171	Sequence 171, Appl
27	2351	100.0	454	US-10-455-822-172	Sequence 172, Appl
28	2351	100.0	454	US-10-455-822-173	Sequence 173, Appl
29	2351	100.0	454	US-10-455-822-174	Sequence 174, Appl
30	2351	100.0	454	US-10-455-822-175	Sequence 175, Appl
31	2351	100.0	454	US-10-455-822-176	Sequence 176, Appl
32	2351	100.0	454	US-10-455-822-177	Sequence 177, Appl
33	2351	100.0	454	US-10-455-822-178	Sequence 178, Appl
34	2351	100.0	454	US-10-455-822-179	Sequence 179, Appl
35	2351	100.0	454	US-10-455-822-180	Sequence 180, Appl
36	2351	100.0	454	US-10-455-822-181	Sequence 181, Appl
37	2351	100.0	454	US-10-455-822-182	Sequence 182, Appl
38	2351	100.0	454	US-10-455-822-183	Sequence 183, Appl
39	2351	100.0	454	US-10-455-822-184	Sequence 184, Appl
40	2351	100.0	454	US-10-455-822-185	Sequence 185, Appl
41	2351	100.0	454	US-10-455-822-186	Sequence 186, Appl
42	2351	100.0	454	US-10-455-822-187	Sequence 187, Appl
43	2351	100.0	454	US-10-455-822-188	Sequence 188, Appl
44	2351	100.0	454	US-10-455-822-189	Sequence 189, Appl
45	2351	100.0	454	US-10-455-822-190	Sequence 190, Appl

ALIGNMENTS

RESULT 1
US-09-888-257A-10
Sequence 10, Application US/09888257A
Publication No. US20030060612A1
GENERAL INFORMATION:
APPLICANT: Goddard, Andrew
APPLICANT: Guney, Austin J.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Polakis, Paul
APPLICANT: Smith, Victoria
APPLICANT: Wood, William I.
APPLICANT: Wu, Thomas D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
TITILE OF INVENTION: TREATMENT OF TUMOR
FILE REFERENCE: P5002R1
CURRENT APPLICATION NUMBER: US/09/888,257A
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: US-60/099,792 NO
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US-60/099,792 NO
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: US-60/103,678 NO
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: US-60/235,451
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US00/04342
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439

Pro 23203
divided 9/24/2000
601235451

PRIOR FILING DATE: 2000-03-30
 PRIOR APPLICATION NUMBER: PCT/US00/23328
 PRIOR FILING DATE: 2000-08-24
 PRIOR APPLICATION NUMBER: PCT/US00/32678
 PRIOR FILING DATE: 2000-12-01
 PRIOR APPLICATION NUMBER: PCT/US01/06520
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: PCT/US01/06666
 PRIOR FILING DATE: 2001-03-01
 NUMBER OF SEQ ID NOS: 10
 SEQ ID NO 10
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-888-257A-10

Query Match 100.0%; Score 2351; DB 10; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6,1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPFASKLITRLIRGCHVTVIGS 60
 DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPFASKLITRLIRGCHVTVIGS 60
 QY 61 RNPKEASEFPFHVVDVTHEDALTKTNIIFVAIHREHTSLMDLRHLLVKGKLLIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHEDALTKTNIIFVAIHREHTSLMDLRHLLVKGKLLIDVSNM 120
 QY 121 RINQYPSNAEYIASLFPDLSLVKGFNVASAWALQGPKDSROYTICSNNOARQOYIE 180
 DB 121 RINQYPSNAEYIASLFPDLSLVKGFNVASAWALQGPKDSROYTICSNNOARQOYIE 180
 QY 181 LARQNLFIPIIDGSSSAREIENLPRLFTLRGFWVAISLATEFFLYSFVRDIHPYA 240
 DB 181 LARQNLFIPIIDGSSSAREIENLPRLFTLRGFWVAISLATEFFLYSFVRDIHPYA 240
 QY 241 RNOQSDFYKIPFETVNTKLPPIVAITLLSLVYLAGLAAAYQYGTGKRRPPEMLWTQ 300
 DB 241 RNOQSDFYKIPFETVNTKLPPIVAITLLSLVYLAGLAAAYQYGTGKRRPPEMLWTQ 300
 QY 301 CRKQGLISFFPAMVAVASLCPMRSEERYFLNMAVOQVHANIENSNWEEVRIEMY 360
 DB 301 CRKQGLISFFPAMVAVASLCPMRSEERYFLNMAVOQVHANIENSNWEEVRIEMY 360
 QY 361 ISFGIMSLGLSLAVTSPSVSNALNMRSEFQISTGLGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAVTSPSVSNALNMRSEFQISTGLGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EYYRFPYTPNFVLAIVLPISIVILDLOLCRYPD 454
 DB 421 EYYRFPYTPNFVLAIVLPISIVILDLOLCRYPD 454

RESULT 2
 US-10-455-822-3

Sequence 3, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Challita-Eld, Pia M.
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P486 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016,24
 CURRENT APPLICATION NUMBER: US/10/455,822
 PRIOR FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-3

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6,1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPFASKLITRLIRGCHVTVIGS 60
 DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPFASKLITRLIRGCHVTVIGS 60
 QY 61 RNPKEASEFPFHVVDVTHEDALTKTNIIFVAIHREHTSLMDLRHLLVKGKLLIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHEDALTKTNIIFVAIHREHTSLMDLRHLLVKGKLLIDVSNM 120
 QY 121 RINQYPSNAEYIASLFPDLSLVKGFNVASAWALQGPKDSROYTICSNNOARQOYIE 180
 DB 121 RINQYPSNAEYIASLFPDLSLVKGFNVASAWALQGPKDSROYTICSNNOARQOYIE 180
 QY 181 LARQNLFIPIIDGSSSAREIENLPRLFTLRGFWVAISLATEFFLYSFVRDIHPYA 240
 DB 181 LARQNLFIPIIDGSSSAREIENLPRLFTLRGFWVAISLATEFFLYSFVRDIHPYA 240
 QY 241 RNOQSDFYKIPFETVNTKLPPIVAITLLSLVYLAGLAAAYQYGTGKRRPPEMLWTQ 300
 DB 241 RNOQSDFYKIPFETVNTKLPPIVAITLLSLVYLAGLAAAYQYGTGKRRPPEMLWTQ 300
 QY 301 CRKQGLISFFPAMVAVASLCPMRSEERYFLNMAVOQVHANIENSNWEEVRIEMY 360
 DB 301 CRKQGLISFFPAMVAVASLCPMRSEERYFLNMAVOQVHANIENSNWEEVRIEMY 360
 QY 361 ISFGIMSLGLSLAVTSPSVSNALNMRSEFQISTGLGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAVTSPSVSNALNMRSEFQISTGLGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EYYRFPYTPNFVLAIVLPISIVILDLOLCRYPD 454
 DB 421 EYYRFPYTPNFVLAIVLPISIVILDLOLCRYPD 454

RESULT 3
 US-10-455-822-7

Sequence 7, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Challita-Eld, Pia M.
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P486 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016,24
 CURRENT APPLICATION NUMBER: US/10/455,822
 PRIOR FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 7
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-7

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6, 1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVYVIGS 60
 DB 1 MESISMWSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVYVIGS 60
 QY 61 RNPKFASFPFHVDVTHEDALTKNIIFVAIHREHYTSLMDRLHVLGKILIDVSNM 120
 DB 61 RNPKFASFPFHVDVTHEDALTKNIIFVAIHREHYTSLMDRLHVLGKILIDVSNM 120
 QY 121 RINQYPSNABYLALSPFDSLIVKGFNVASAMALQCPKASROYVICSNNIQARQVIE 180
 DB 121 RINQYPSNABYLALSPFDSLIVKGFNVASAMALQCPKASROYVICSNNIQARQVIE 180
 QY 121 RINQYPSNABYLALSPFDSLIVKGFNVASAMALQCPKASROYVICSNNIQARQVIE 180
 DB 121 RINQYPSNABYLALSPFDSLIVKGFNVASAMALQCPKASROYVICSNNIQARQVIE 180
 QY 181 LARQNFIPIDGLSISAREIENLPRLFTLMRGPVVAISLAFPFVLSFVRDVHPYA 240
 DB 181 LARQNFIPIDGLSISAREIENLPRLFTLMRGPVVAISLAFPFVLSFVRDVHPYA 240
 QY 181 LARQNFIPIDGLSISAREIENLPRLFTLMRGPVVAISLAFPFVLSFVRDVHPYA 240
 DB 181 LARQNFIPIDGLSISAREIENLPRLFTLMRGPVVAISLAFPFVLSFVRDVHPYA 240
 QY 241 RNOQSDFYKIPRIEIVNKTLPVATITLSLVYLAGLAAAYOLYGTXYRRPPLMTWQ 300
 DB 241 RNOQSDFYKIPRIEIVNKTLPVATITLSLVYLAGLAAAYOLYGTXYRRPPLMTWQ 300
 QY 241 RNOQSDFYKIPRIEIVNKTLPVATITLSLVYLAGLAAAYOLYGTXYRRPPLMTWQ 300
 DB 241 RNOQSDFYKIPRIEIVNKTLPVATITLSLVYLAGLAAAYOLYGTXYRRPPLMTWQ 300
 QY 301 CRKQGLISFFPAMVAVSYCLPWRSEERYLFLNMAVQVHANIEISMNEEYWRILEMY 360
 DB 301 CRKQGLISFFPAMVAVSYCLPWRSEERYLFLNMAVQVHANIEISMNEEYWRILEMY 360
 QY 301 CRKQGLISFFPAMVAVSYCLPWRSEERYLFLNMAVQVHANIEISMNEEYWRILEMY 360
 DB 301 CRKQGLISFFPAMVAVSYCLPWRSEERYLFLNMAVQVHANIEISMNEEYWRILEMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTLGYVALLISTPHVLIYGMKQAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTLGYVALLISTPHVLIYGMKQAFE 420
 QY 421 EBYRFTYTPNPFVLAALVPSIVIIDLLQLCRYPD 454
 DB 421 EBYRFTYTPNPFVLAALVPSIVIIDLLQLCRYPD 454

RESULT 4
 US-10-455-822-9
 Sequence 9, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Mangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Challita-Bid, Pia M.
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 PRIOR FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 9
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-9

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6, 1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVYVIGS 60
 DB 1 MESISMWSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVYVIGS 60
 QY 61 RNPKFASFPFHVDVTHEDALTKNIIFVAIHREHYTSLMDRLHVLGKILIDVSNM 120
 DB 61 RNPKFASFPFHVDVTHEDALTKNIIFVAIHREHYTSLMDRLHVLGKILIDVSNM 120
 QY 121 RINQYPSNABYLALSPFDSLIVKGFNVASAMALQCPKASROYVICSNNIQARQVIE 180
 DB 121 RINQYPSNABYLALSPFDSLIVKGFNVASAMALQCPKASROYVICSNNIQARQVIE 180
 QY 121 RINQYPSNABYLALSPFDSLIVKGFNVASAMALQCPKASROYVICSNNIQARQVIE 180
 DB 121 RINQYPSNABYLALSPFDSLIVKGFNVASAMALQCPKASROYVICSNNIQARQVIE 180
 QY 181 LARQNFIPIDGLSISAREIENLPRLFTLMRGPVVAISLAFPFVLSFVRDVHPYA 240
 DB 181 LARQNFIPIDGLSISAREIENLPRLFTLMRGPVVAISLAFPFVLSFVRDVHPYA 240
 QY 181 LARQNFIPIDGLSISAREIENLPRLFTLMRGPVVAISLAFPFVLSFVRDVHPYA 240
 DB 181 LARQNFIPIDGLSISAREIENLPRLFTLMRGPVVAISLAFPFVLSFVRDVHPYA 240
 QY 241 RNOQSDFYKIPRIEIVNKTLPVATITLSLVYLAGLAAAYOLYGTXYRRPPLMTWQ 300
 DB 241 RNOQSDFYKIPRIEIVNKTLPVATITLSLVYLAGLAAAYOLYGTXYRRPPLMTWQ 300
 QY 241 RNOQSDFYKIPRIEIVNKTLPVATITLSLVYLAGLAAAYOLYGTXYRRPPLMTWQ 300
 DB 241 RNOQSDFYKIPRIEIVNKTLPVATITLSLVYLAGLAAAYOLYGTXYRRPPLMTWQ 300
 QY 301 CRKQGLISFFPAMVAVSYCLPWRSEERYLFLNMAVQVHANIEISMNEEYWRILEMY 360
 DB 301 CRKQGLISFFPAMVAVSYCLPWRSEERYLFLNMAVQVHANIEISMNEEYWRILEMY 360
 QY 301 CRKQGLISFFPAMVAVSYCLPWRSEERYLFLNMAVQVHANIEISMNEEYWRILEMY 360
 DB 301 CRKQGLISFFPAMVAVSYCLPWRSEERYLFLNMAVQVHANIEISMNEEYWRILEMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTLGYVALLISTPHVLIYGMKQAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTLGYVALLISTPHVLIYGMKQAFE 420
 QY 421 EBYRFTYTPNPFVLAALVPSIVIIDLLQLCRYPD 454
 DB 421 EBYRFTYTPNPFVLAALVPSIVIIDLLQLCRYPD 454

RESULT 5
 US-10-455-822-19
 Sequence 19, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Mangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Challita-Bid, Pia M.
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 PRIOR FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 19
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-19

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6,1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSTCLPNGINGIKDKAKTVGVIGSGDPKSLTIRLRIGYHVYIGS 60
 DB 1 MESISMGSPKSLSTCLPNGINGIKDKAKTVGVIGSGDPKSLTIRLRIGYHVYIGS 60
 QY 61 RNPKESEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKESEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQVPSNAEYLAIFPDSLIVKGFNVVSAMALQGPDKASROYICSNNOQAQOYIE 180
 DB 121 RINQVPSNAEYLAIFPDSLIVKGFNVVSAMALQGPDKASROYICSNNOQAQOYIE 180
 QY 181 LARQNFIDIDGSLSSAREIENLPLRFTLMRGVVAVIAISLATEFFLYSPRVDIHRYA 240
 DB 181 LARQNFIDIDGSLSSAREIENLPLRFTLMRGVVAVIAISLATEFFLYSPRVDIHRYA 240
 QY 241 RNOQSDFKIPLEIYNKTLPIVAITLLSLVYAGLLAAVQYVGTKRRPPLMETWQ 300
 DB 241 RNOQSDFKIPLEIYNKTLPIVAITLLSLVYAGLLAAVQYVGTKRRPPLMETWQ 300
 QY 301 CRKQGLSFFPAMVAVSLCLPMRSERYFLNMAVOQHANIENSNNEEVRIMY 360
 DB 301 CRKQGLSFFPAMVAVSLCLPMRSERYFLNMAVOQHANIENSNNEEVRIMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTIGYALLISTFHVLIYKKAPE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTIGYALLISTFHVLIYKKAPE 420
 QY 421 EBYRFPYTPNFVLAIVLPSIVILDLLQCRYPD 454
 DB 421 EBYRFPYTPNFVLAIVLPSIVILDLLQCRYPD 454

RESULT 6

US-10-455-822-21
 Sequence 21, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chalita-Eld, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 21
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-21

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6,1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSTCLPNGINGIKDKAKTVGVIGSGDPKSLTIRLRIGYHVYIGS 60
 DB 1 MESISMGSPKSLSTCLPNGINGIKDKAKTVGVIGSGDPKSLTIRLRIGYHVYIGS 60
 QY 61 RNPKESEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKESEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQVPSNAEYLAIFPDSLIVKGFNVVSAMALQGPDKASROYICSNNOQAQOYIE 180
 DB 121 RINQVPSNAEYLAIFPDSLIVKGFNVVSAMALQGPDKASROYICSNNOQAQOYIE 180
 QY 181 LARQNFIDIDGSLSSAREIENLPLRFTLMRGVVAVIAISLATEFFLYSPRVDIHRYA 240
 DB 181 LARQNFIDIDGSLSSAREIENLPLRFTLMRGVVAVIAISLATEFFLYSPRVDIHRYA 240
 QY 241 RNOQSDFKIPLEIYNKTLPIVAITLLSLVYAGLLAAVQYVGTKRRPPLMETWQ 300
 DB 241 RNOQSDFKIPLEIYNKTLPIVAITLLSLVYAGLLAAVQYVGTKRRPPLMETWQ 300
 QY 301 CRKQGLSFFPAMVAVSLCLPMRSERYFLNMAVOQHANIENSNNEEVRIMY 360
 DB 301 CRKQGLSFFPAMVAVSLCLPMRSERYFLNMAVOQHANIENSNNEEVRIMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTIGYALLISTFHVLIYKKAPE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTIGYALLISTFHVLIYKKAPE 420
 QY 421 EBYRFPYTPNFVLAIVLPSIVILDLLQCRYPD 454
 DB 421 EBYRFPYTPNFVLAIVLPSIVILDLLQCRYPD 454

RESULT 7

US-10-455-822-23
 Sequence 23, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chalita-Eld, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-23

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6,1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMSPKSLSECTCPNGINGIKDARKTVGVIGSGDPKSLTIRIRCGYHVYIGS 60
DB 1 MESISMSPKSLSECTCPNGINGIKDARKTVGVIGSGDPKSLTIRIRCGYHVYIGS 60
QY 61 RNPKFASFPFPHVVDVTHEDALTKTNII FVAIHREHYTSLMDRLHLLVGKILIDVSNM 120
DB 61 RNPKFASFPFPHVVDVTHEDALTKTNII FVAIHREHYTSLMDRLHLLVGKILIDVSNM 120
QY 121 RINQYPSNAEYLAISFPDSLIVKGFNVASAMALQGPDKASROYICSNNOIQAQOYIE 180
DB 121 RINQYPSNAEYLAISFPDSLIVKGFNVASAMALQGPDKASROYICSNNOIQAQOYIE 180
QY 121 RINQYPSNAEYLAISFPDSLIVKGFNVASAMALQGPDKASROYICSNNOIQAQOYIE 180
DB 121 RINQYPSNAEYLAISFPDSLIVKGFNVASAMALQGPDKASROYICSNNOIQAQOYIE 180
QY 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLWRGVPVVAISLATEFFLYSPRVDIHPYA 240
DB 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLWRGVPVVAISLATEFFLYSPRVDIHPYA 240
QY 241 RNOQSDFKYKPIEIVNKTLPVATITLSLVYAGLAAAYQLYGTGKRRPPEMLQ 300
DB 241 RNOQSDFKYKPIEIVNKTLPVATITLSLVYAGLAAAYQLYGTGKRRPPEMLQ 300
QY 241 RNOQSDFKYKPIEIVNKTLPVATITLSLVYAGLAAAYQLYGTGKRRPPEMLQ 300
DB 241 RNOQSDFKYKPIEIVNKTLPVATITLSLVYAGLAAAYQLYGTGKRRPPEMLQ 300
QY 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANINENSNNEEVRIMY 360
DB 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANINENSNNEEVRIMY 360
QY 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANINENSNNEEVRIMY 360
DB 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANINENSNNEEVRIMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTIGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTIGYVALLISTFHVLIYGMKRAFE 420
QY 421 EYYRFYTPPNFVALVLPISIVILDLOLCRYPD 454
DB 421 EYYRFYTPPNFVALVLPISIVILDLOLCRYPD 454

RESULT 8
US-10-455-822-25

Sequence 25, Application US/10455822
Publication No. US20040048798A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Raitano, Arthur B.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovits, Aya
APPLICANT: Chailita-Bid, Pia M.
APPLICANT: Farris, Mary
TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
TITLE OF INVENTION: Entitled 98486 Useful in Treatment and Detection of Cancer
FILE REFERENCE: 51158-20016.24
CURRENT APPLICATION NUMBER: US/10/455,822
CURRENT FILING DATE: 2003-06-04
PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US60/370,387
PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-25

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6,1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMSPKSLSECTCPNGINGIKDARKTVGVIGSGDPKSLTIRIRCGYHVYIGS 60
DB 1 MESISMSPKSLSECTCPNGINGIKDARKTVGVIGSGDPKSLTIRIRCGYHVYIGS 60
QY 61 RNPKFASFPFPHVVDVTHEDALTKTNII FVAIHREHYTSLMDRLHLLVGKILIDVSNM 120
DB 61 RNPKFASFPFPHVVDVTHEDALTKTNII FVAIHREHYTSLMDRLHLLVGKILIDVSNM 120
QY 121 RINQYPSNAEYLAISFPDSLIVKGFNVASAMALQGPDKASROYICSNNOIQAQOYIE 180
DB 121 RINQYPSNAEYLAISFPDSLIVKGFNVASAMALQGPDKASROYICSNNOIQAQOYIE 180
QY 121 RINQYPSNAEYLAISFPDSLIVKGFNVASAMALQGPDKASROYICSNNOIQAQOYIE 180
DB 121 RINQYPSNAEYLAISFPDSLIVKGFNVASAMALQGPDKASROYICSNNOIQAQOYIE 180
QY 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLWRGVPVVAISLATEFFLYSPRVDIHPYA 240
DB 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLWRGVPVVAISLATEFFLYSPRVDIHPYA 240
QY 241 RNOQSDFKYKPIEIVNKTLPVATITLSLVYAGLAAAYQLYGTGKRRPPEMLQ 300
DB 241 RNOQSDFKYKPIEIVNKTLPVATITLSLVYAGLAAAYQLYGTGKRRPPEMLQ 300
QY 241 RNOQSDFKYKPIEIVNKTLPVATITLSLVYAGLAAAYQLYGTGKRRPPEMLQ 300
DB 241 RNOQSDFKYKPIEIVNKTLPVATITLSLVYAGLAAAYQLYGTGKRRPPEMLQ 300
QY 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANINENSNNEEVRIMY 360
DB 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANINENSNNEEVRIMY 360
QY 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANINENSNNEEVRIMY 360
DB 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANINENSNNEEVRIMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTIGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTIGYVALLISTFHVLIYGMKRAFE 420
QY 421 EYYRFYTPPNFVALVLPISIVILDLOLCRYPD 454
DB 421 EYYRFYTPPNFVALVLPISIVILDLOLCRYPD 454

RESULT 9
US-10-455-822-31

Sequence 31, Application US/10455822
Publication No. US20040048798A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Raitano, Arthur B.
APPLICANT: Ge, Wangmao
APPLICANT: Jakobovits, Aya
APPLICANT: Chailita-Bid, Pia M.
APPLICANT: Farris, Mary
TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
TITLE OF INVENTION: Entitled 98486 Useful in Treatment and Detection of Cancer
FILE REFERENCE: 51158-20016.24
CURRENT APPLICATION NUMBER: US/10/455,822
CURRENT FILING DATE: 2003-06-04
PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US60/370,387
PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 31
 LENGTH: 454
 TYPE: PR
 ORGANISM: Homo sapien
 US-10-455-822-31

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSTCLPNGINGIKDKAKTYGVIGSGDPKSLTIRLRICGHHVIGS 60
 DB 1 MESISMGSPKSLSTCLPNGINGIKDKAKTYGVIGSGDPKSLTIRLRICGHHVIGS 60
 QY 61 RNPKASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQYENAEYLAFLPDSLLVKGPNVSAMALQGPKDSROYICSNNIQARQVIE 180
 DB 121 RINQYENAEYLAFLPDSLLVKGPNVSAMALQGPKDSROYICSNNIQARQVIE 180
 QY 121 RINQYENAEYLAFLPDSLLVKGPNVSAMALQGPKDSROYICSNNIQARQVIE 180
 DB 121 RINQYENAEYLAFLPDSLLVKGPNVSAMALQGPKDSROYICSNNIQARQVIE 180
 QY 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGVVAISLATFFLYSFVRDIHPYA 240
 DB 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGVVAISLATFFLYSFVRDIHPYA 240
 QY 241 RNOQSDFYKIPLEIVNKTLPVIAITLLSVYLAGLAAAYOLYGTXYRRPFWLETLQ 300
 DB 241 RNOQSDFYKIPLEIVNKTLPVIAITLLSVYLAGLAAAYOLYGTXYRRPFWLETLQ 300
 QY 301 CRKQGLISFFPAMVAVASLCLPMRSERYLFLMAVQOVANINENSMNEEVRILEMY 360
 DB 301 CRKQGLISFFPAMVAVASLCLPMRSERYLFLMAVQOVANINENSMNEEVRILEMY 360
 QY 361 ISFGISGLISLAVTSIPSVSNALNMRREFSFIQSTIGYVALLISTFHVLIYWKRAFE 420
 DB 361 ISFGISGLISLAVTSIPSVSNALNMRREFSFIQSTIGYVALLISTFHVLIYWKRAFE 420
 QY 421 EBYRFTYTPNFVLAIVLPSIYILDLOLCRYPD 454
 DB 421 EBYRFTYTPNFVLAIVLPSIYILDLOLCRYPD 454

RESULT 10

US-10-455-822-33
 Sequence 33, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Eid, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P486 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 33
 LENGTH: 454
 TYPE: PR
 ORGANISM: Homo sapien
 US-10-455-822-33

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSTCLPNGINGIKDKAKTYGVIGSGDPKSLTIRLRICGHHVIGS 60
 DB 1 MESISMGSPKSLSTCLPNGINGIKDKAKTYGVIGSGDPKSLTIRLRICGHHVIGS 60
 QY 61 RNPKASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKASEFFPHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQYENAEYLAFLPDSLLVKGPNVSAMALQGPKDSROYICSNNIQARQVIE 180
 DB 121 RINQYENAEYLAFLPDSLLVKGPNVSAMALQGPKDSROYICSNNIQARQVIE 180
 QY 121 RINQYENAEYLAFLPDSLLVKGPNVSAMALQGPKDSROYICSNNIQARQVIE 180
 DB 121 RINQYENAEYLAFLPDSLLVKGPNVSAMALQGPKDSROYICSNNIQARQVIE 180
 QY 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGVVAISLATFFLYSFVRDIHPYA 240
 DB 181 LARQNFIPIDIGLSASAREIENLPLRLFTLMRGVVAISLATFFLYSFVRDIHPYA 240
 QY 241 RNOQSDFYKIPLEIVNKTLPVIAITLLSVYLAGLAAAYOLYGTXYRRPFWLETLQ 300
 DB 241 RNOQSDFYKIPLEIVNKTLPVIAITLLSVYLAGLAAAYOLYGTXYRRPFWLETLQ 300
 QY 301 CRKQGLISFFPAMVAVASLCLPMRSERYLFLMAVQOVANINENSMNEEVRILEMY 360
 DB 301 CRKQGLISFFPAMVAVASLCLPMRSERYLFLMAVQOVANINENSMNEEVRILEMY 360
 QY 361 ISFGISGLISLAVTSIPSVSNALNMRREFSFIQSTIGYVALLISTFHVLIYWKRAFE 420
 DB 361 ISFGISGLISLAVTSIPSVSNALNMRREFSFIQSTIGYVALLISTFHVLIYWKRAFE 420
 QY 421 EBYRFTYTPNFVLAIVLPSIYILDLOLCRYPD 454
 DB 421 EBYRFTYTPNFVLAIVLPSIYILDLOLCRYPD 454

RESULT 11

US-10-455-822-35
 Sequence 35, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Eid, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P486 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 35
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-35

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
 DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
 QY 61 RNPKPASEFPFHVVDVTHEDALTKNTIIFVAIHEHYTSLMDLRHLVGLKILIDVSNM 120
 DB 61 RNPKPASEFPFHVVDVTHEDALTKNTIIFVAIHEHYTSLMDLRHLVGLKILIDVSNM 120
 QY 121 RINQVPSNAEYLAIFPDSLIVKGFNVSAWALQGPDKASROYICSNNOARQOYIE 180
 DB 121 RINQVPSNAEYLAIFPDSLIVKGFNVSAWALQGPDKASROYICSNNOARQOYIE 180
 QY 121 RINQVPSNAEYLAIFPDSLIVKGFNVSAWALQGPDKASROYICSNNOARQOYIE 180
 DB 121 RINQVPSNAEYLAIFPDSLIVKGFNVSAWALQGPDKASROYICSNNOARQOYIE 180
 QY 181 LARQNLFIPIIDGSLSSAREIENLPLRLFTLMRGVVAISLATEFFLYSPFVDVHHPYA 240
 DB 181 LARQNLFIPIIDGSLSSAREIENLPLRLFTLMRGVVAISLATEFFLYSPFVDVHHPYA 240
 QY 241 RNOQSDFYKIPLEIYNKTLPIVAITLISLVYLAGLAAAYQLYGTXRRPPELWLMQ 300
 DB 241 RNOQSDFYKIPLEIYNKTLPIVAITLISLVYLAGLAAAYQLYGTXRRPPELWLMQ 300
 QY 241 RNOQSDFYKIPLEIYNKTLPIVAITLISLVYLAGLAAAYQLYGTXRRPPELWLMQ 300
 DB 241 RNOQSDFYKIPLEIYNKTLPIVAITLISLVYLAGLAAAYQLYGTXRRPPELWLMQ 300
 QY 301 CRKQGLSLFFPAMVAVASLCLPWRRSERYFLNMAVQVHANENSWNEEVRIMY 360
 DB 301 CRKQGLSLFFPAMVAVASLCLPWRRSERYFLNMAVQVHANENSWNEEVRIMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFISOSTIGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFISOSTIGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EBYRFFYTPPNFVALVLPISIVILDLQLCRYPD 454
 DB 421 EBYRFFYTPPNFVALVLPISIVILDLQLCRYPD 454

RESULT 12

US-10-455-822-37
 Sequence 37, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chaillita-Eld, Pia M.
 APPLICANT: Paris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 9894B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 37
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-37

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
 DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCVHVG 60
 QY 61 RNPKPASEFPFHVVDVTHEDALTKNTIIFVAIHEHYTSLMDLRHLVGLKILIDVSNM 120
 DB 61 RNPKPASEFPFHVVDVTHEDALTKNTIIFVAIHEHYTSLMDLRHLVGLKILIDVSNM 120
 QY 121 RINQVPSNAEYLAIFPDSLIVKGFNVSAWALQGPDKASROYICSNNOARQOYIE 180
 DB 121 RINQVPSNAEYLAIFPDSLIVKGFNVSAWALQGPDKASROYICSNNOARQOYIE 180
 QY 121 RINQVPSNAEYLAIFPDSLIVKGFNVSAWALQGPDKASROYICSNNOARQOYIE 180
 DB 121 RINQVPSNAEYLAIFPDSLIVKGFNVSAWALQGPDKASROYICSNNOARQOYIE 180
 QY 181 LARQNLFIPIIDGSLSSAREIENLPLRLFTLMRGVVAISLATEFFLYSPFVDVHHPYA 240
 DB 181 LARQNLFIPIIDGSLSSAREIENLPLRLFTLMRGVVAISLATEFFLYSPFVDVHHPYA 240
 QY 241 RNOQSDFYKIPLEIYNKTLPIVAITLISLVYLAGLAAAYQLYGTXRRPPELWLMQ 300
 DB 241 RNOQSDFYKIPLEIYNKTLPIVAITLISLVYLAGLAAAYQLYGTXRRPPELWLMQ 300
 QY 241 RNOQSDFYKIPLEIYNKTLPIVAITLISLVYLAGLAAAYQLYGTXRRPPELWLMQ 300
 DB 241 RNOQSDFYKIPLEIYNKTLPIVAITLISLVYLAGLAAAYQLYGTXRRPPELWLMQ 300
 QY 301 CRKQGLSLFFPAMVAVASLCLPWRRSERYFLNMAVQVHANENSWNEEVRIMY 360
 DB 301 CRKQGLSLFFPAMVAVASLCLPWRRSERYFLNMAVQVHANENSWNEEVRIMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFISOSTIGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFISOSTIGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EBYRFFYTPPNFVALVLPISIVILDLQLCRYPD 454
 DB 421 EBYRFFYTPPNFVALVLPISIVILDLQLCRYPD 454

RESULT 13

US-10-455-822-39
 Sequence 39, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chaillita-Eld, Pia M.
 APPLICANT: Paris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 9894B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 39
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-39

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCPLPNCINGIKDARKYTVGVISGDFAKSLITRLIRCGYHVIGS 60
 DB 1 MESISMWSPKSLSETCPLPNCINGIKDARKYTVGVISGDFAKSLITRLIRCGYHVIGS 60
 QY 61 RNPKEASEFPFHVDVTHEDALTKNIIFFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKEASEFPFHVDVTHEDALTKNIIFFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQPEESNAEYLASLPFDSILVGFNVVSAMALQLPKASRQVYICSNNIQARQVIE 180
 DB 121 RINQPEESNAEYLASLPFDSILVGFNVVSAMALQLPKASRQVYICSNNIQARQVIE 180
 QY 121 RINQPEESNAEYLASLPFDSILVGFNVVSAMALQLPKASRQVYICSNNIQARQVIE 180
 DB 121 RINQPEESNAEYLASLPFDSILVGFNVVSAMALQLPKASRQVYICSNNIQARQVIE 180
 QY 181 LARQNLFPIDIGLSAREIENPLRLFTLMRGPVVAISLAFPELYSFVRVHPYA 240
 DB 181 LARQNLFPIDIGLSAREIENPLRLFTLMRGPVVAISLAFPELYSFVRVHPYA 240
 QY 241 RNOQSDFYKPIEIVNKTLPVATITLSLVYLAGLAAAYQLYGTXYRRPFWLBTWQ 300
 DB 241 RNOQSDFYKPIEIVNKTLPVATITLSLVYLAGLAAAYQLYGTXYRRPFWLBTWQ 300
 QY 301 CRKQGLISFFPAMVAVSYCLPMRSEERYLFNMAVQVHANIEENMEEEVWRIEMV 360
 DB 301 CRKQGLISFFPAMVAVSYCLPMRSEERYLFNMAVQVHANIEENMEEEVWRIEMV 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFTIOSTIGYVALLISTFHVLIYGMKCAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFTIOSTIGYVALLISTFHVLIYGMKCAFE 420
 QY 421 EEEYRFYTPPNFVALVLPISVILDLQLCRYPD 454
 DB 421 EEEYRFYTPPNFVALVLPISVILDLQLCRYPD 454

RESULT 14
 US-10-455-822-78

Sequence 78, Application US/10455822
 Publication No. US20040048798A1

GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Bid, Pia M.
 APPLICANT: Paris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 989486 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 78
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-78

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCPLPNCINGIKDARKYTVGVISGDFAKSLITRLIRCGYHVIGS 60
 DB 1 MESISMWSPKSLSETCPLPNCINGIKDARKYTVGVISGDFAKSLITRLIRCGYHVIGS 60
 QY 61 RNPKEASEFPFHVDVTHEDALTKNIIFFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKEASEFPFHVDVTHEDALTKNIIFFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQPEESNAEYLASLPFDSILVGFNVVSAMALQLPKASRQVYICSNNIQARQVIE 180
 DB 121 RINQPEESNAEYLASLPFDSILVGFNVVSAMALQLPKASRQVYICSNNIQARQVIE 180
 QY 121 RINQPEESNAEYLASLPFDSILVGFNVVSAMALQLPKASRQVYICSNNIQARQVIE 180
 DB 121 RINQPEESNAEYLASLPFDSILVGFNVVSAMALQLPKASRQVYICSNNIQARQVIE 180
 QY 181 LARQNLFPIDIGLSAREIENPLRLFTLMRGPVVAISLAFPELYSFVRVHPYA 240
 DB 181 LARQNLFPIDIGLSAREIENPLRLFTLMRGPVVAISLAFPELYSFVRVHPYA 240
 QY 241 RNOQSDFYKPIEIVNKTLPVATITLSLVYLAGLAAAYQLYGTXYRRPFWLBTWQ 300
 DB 241 RNOQSDFYKPIEIVNKTLPVATITLSLVYLAGLAAAYQLYGTXYRRPFWLBTWQ 300
 QY 301 CRKQGLISFFPAMVAVSYCLPMRSEERYLFNMAVQVHANIEENMEEEVWRIEMV 360
 DB 301 CRKQGLISFFPAMVAVSYCLPMRSEERYLFNMAVQVHANIEENMEEEVWRIEMV 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFTIOSTIGYVALLISTFHVLIYGMKCAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFTIOSTIGYVALLISTFHVLIYGMKCAFE 420
 QY 421 EEEYRFYTPPNFVALVLPISVILDLQLCRYPD 454
 DB 421 EEEYRFYTPPNFVALVLPISVILDLQLCRYPD 454

RESULT 15
 US-10-455-822-121

Sequence 121, Application US/10455822
 Publication No. US20040048798A1

GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Bid, Pia M.
 APPLICANT: Paris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 989486 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 121
LENGTH: 454
TYPE: PRT
ORGANISM: Homo sapien
US-10-455-822-121

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMNGSPKSLSETCIPNGINGIKDARKTVGVIGSGDPAKSLTIRLRGCVHVVIGS 60
DB 1 MESISMNGSPKSLSETCIPNGINGIKDARKTVGVIGSGDPAKSLTIRLRGCVHVVIGS 60
QY 61 RNPFASEFFPHVVDVTHEDALTKTNIIFAIHREHYTSLMDLRHLLVGKILIDVSNM 120
DB 61 RNPFASEFFPHVVDVTHEDALTKTNIIFAIHREHYTSLMDLRHLLVGKILIDVSNM 120
QY 121 RINQYBESNAEYLASLPDLSIIVKGFNVVSAMALQGPKASROYICSNNIQARQVIE 180
DB 121 RINQYBESNAEYLASLPDLSIIVKGFNVVSAMALQGPKASROYICSNNIQARQVIE 180
QY 181 LARQNFIPIDIGSLISAREIENLPLRFTLMRGPVVAISLATFFLYSFVDVIRHYA 240
DB 181 LARQNFIPIDIGSLISAREIENLPLRFTLMRGPVVAISLATFFLYSFVDVIRHYA 240
QY 241 RNQOSDFKIPIEIVNKTLPVATITLSVYLAGLAAAYQLYGCKRRPPLWETWLO 300
DB 241 RNQOSDFKIPIEIVNKTLPVATITLSVYLAGLAAAYQLYGCKRRPPLWETWLO 300
QY 301 CRKQGLSPFPAMVHVAISCLPWRSEERYLFLNMAVQVHANIEINSWNEEYVRIEMY 360
DB 301 CRKQGLSPFPAMVHVAISCLPWRSEERYLFLNMAVQVHANIEINSWNEEYVRIEMY 360
QY 361 ISFGIMSLGLISLAVTSIPSVSNALNKRSEFIQSTIGYVALLISTHVLIIYWKRAFE 420
DB 361 ISFGIMSLGLISLAVTSIPSVSNALNKRSEFIQSTIGYVALLISTHVLIIYWKRAFE 420
QY 421 EBYRYFTPPNFVALVLPISVILDLQLCRYPD 454
DB 421 EBYRYFTPPNFVALVLPISVILDLQLCRYPD 454

Search completed: April 16, 2004, 12:46:36
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:38:02 ; Search time 22 Seconds
(without alignments) 1065.372 Million cell updates/sec

Title: US-09-455-486-6
Sequence: 1 MESTSMWGSKSLSETCLPN.....ALVLPISVLDLQLCRYPD 454

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCUTS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	901	38.3	173	4 US-09-323-873A-8	Sequence 8, Appli
2	736	31.3	141	3 US-09-083-521-1	Sequence 1, Appli
3	717	30.5	339	4 US-09-323-873A-2	Sequence 2, Appli
4	717	30.5	339	4 US-09-685-166A-879	Sequence 879, App
5	144.5	6.1	227	4 US-09-655-270A-15	Sequence 15, Appli
6	144.5	6.1	227	4 US-09-651-941-17	Sequence 17, Appli
7	144.5	6.1	227	4 US-09-955-597-17	Sequence 17, Appli
8	107.5	4.6	695	1 US-08-487-886-2	Sequence 2, Appli
9	107.5	4.6	695	1 US-08-482-855-2	Sequence 2, Appli
10	107.5	4.6	695	1 US-08-474-986-2	Sequence 2, Appli
11	107.5	4.6	695	1 US-08-474-986-2	Sequence 2, Appli
12	102.5	4.4	940	4 US-09-323-873A-20	Sequence 20, Appli
13	101	4.3	476	4 US-09-316-083-3	Sequence 3, Appli
14	101	4.3	476	4 US-09-323-873A-2	Sequence 2, Appli
15	101	4.3	476	4 US-09-323-873A-2	Sequence 2, Appli
16	99	4.2	365	4 US-09-170-496D-118	Sequence 118, App
17	99	4.2	365	4 US-09-170-496D-118	Sequence 118, App
18	99	4.2	365	4 US-09-364-425B-27	Sequence 27, App
19	99	4.2	365	4 US-08-724-974A-2	Sequence 2, Appli
20	97.5	4.1	390	3 US-08-460-576-2	Sequence 2, Appli
21	97	4.1	682	3 US-07-757-342D-6	Sequence 6, Appli
22	97	4.1	682	3 US-07-757-342D-6	Sequence 6, Appli
23	96	4.1	289	4 US-09-503-456-61	Sequence 61, Appli
24	96	4.1	1309	4 US-09-975-413A-10	Sequence 10, Appli
25	94.5	4.0	288	4 US-09-107-532A-6009	Sequence 6009, Ap
26	94.5	4.0	288	4 US-09-540-236-2922	Sequence 2922, Ap
27	94.5	4.0	345	4 US-09-489-039A-10740	Sequence 10740, A

ALIGNMENTS

28	94	4.0	284	4 US-09-903-456-83	Sequence 83, Appli
29	94	4.0	292	4 US-09-903-456-35	Sequence 35, Appli
30	94	4.0	293	4 US-09-903-456-45	Sequence 45, Appli
31	93	4.0	365	2 US-08-724-974A-3	Sequence 3, Appli
32	93	4.0	498	4 US-09-107-532A-7077	Sequence 7077, Ap
33	92.5	3.9	405	4 US-09-489-039A-9411	Sequence 9411, Ap
34	91.5	3.9	407	4 US-09-328-352-5605	Sequence 5605, Ap
35	91.5	3.9	407	4 US-08-985-343-2	Sequence 2, Appli
36	91	3.9	211	4 US-09-252-991A-23822	Sequence 23822, A
37	91	3.9	440	4 US-09-634-238-281	Sequence 281, App
38	91	3.9	550	4 US-09-543-681A-6026	Sequence 6026, Ap
39	91	3.9	808	4 US-09-134-001C-3105	Sequence 3105, Ap
40	90.5	3.8	340	4 US-09-543-681A-7503	Sequence 7503, Ap
41	90.5	3.8	388	4 US-09-107-532A-6381	Sequence 6381, Ap
42	90	3.8	500	4 US-09-178-053B-26	Sequence 26, Appli
43	90	3.8	500	4 US-09-795-927-7	Sequence 7, Appli
44	89.5	3.8	495	1 US-07-841-997A-2	Sequence 2, Appli
45	89.5	3.8	495	1 US-08-290-301-2	Sequence 2, Appli

RESULT 1
US-09-323-873A-8
Sequence 8, Application US/09323873A
Patent No. 6328503
GENERAL INFORMATION:

APPLICANT: Daniel E. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Kahon Leong
APPLICANT: Arthur B. Raitano
APPLICANT: Douglas C. Saffran
APPLICANT: Steve Chappell Mitchell
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
FILE REFERENCE: 129.160502
CURRENT APPLICATION NUMBER: US/09/323,873A
CURRENT FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087,520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 173
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-323-873A-8

Query Match 38.3%; Score 901; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.4e-86; Indels 0; Gaps 0;
Matches 173; Conservative 0; Mismatches 0;

QY	246	DFKIPRIETNKLPVATITLSLVYLAGLAAQYLYGKRRPPMLTQLCKKQV	305
DB	1	DFYKIPRIETNKLPVATITLSLVYLAGLAAQYLYGKRRPPMLTQLCKKQV	60
QY	306	GLISFPFAMVAVSLCLPFRSEERYFLNMAVQOVHANIENSNNEEYWRIMYSFGI	365
DB	61	GLISFPFAMVAVSLCLPFRSEERYFLNMAVQOVHANIENSNNEEYWRIMYSFGI	120
QY	366	MSGLSLAVTSIPSVSNLNNRBEFIOSTGYVALISTFHYLIYWKRA	418
DB	121	MSGLSLAVTSIPSVSNLNNRBEFIOSTGYVALISTFHYLIYWKRA	173

RESULT 2
US-09-083-521-1
Sequence 1, Application US/09083521
Patent No. 6048970
GENERAL INFORMATION:

```

APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,521
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0527 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTUT10
CLONE: 1691243
US-09-083-521-1

Query Match          31.3%; Score 736; DB 3; Length 141;
Best Local Similarity 100.0%; Pred. No. 2.2e-69;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 MHWVAISLCIPMRSEKYLFLNMAVOOYHANIENSMNEEVRIMYISFGIMSLGLSL 373
DB 1 MHWVAISLCIPMRSEKYLFLNMAVOOYHANIENSMNEEVRIMYISFGIMSLGLSL 60
QY 374 LAVTSPVSNALNMEFEFSFIQSTLGYVALLSTFHVLIYGWKRAFEERYRFTPPNFV 433
DB 61 LAVTSPVSNALNMEFEFSFIQSTLGYVALLSTFHVLIYGWKRAFEERYRFTPPNFV 120
QY 434 LALVPSIVTLDLQCRYPD 454
DB 121 LALVPSIVTLDLQCRYPD 141

RESULT 3
US-09-323-873A-2
Sequence 2, Application US/09323873A
Patent No. 6328503
GENERAL INFORMATION:
APPLICANT: Daniel E. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Kahen Leong
APPLICANT: Arthur B. Raitano
APPLICANT: Douglas C. Saffran
APPLICANT: Steve Chappeil Mitchell
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
FILE REFERENCE: 129.16US02
CURRENT APPLICATION NUMBER: US/09/323,873A
CURRENT FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087,520

```

```

PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 339
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-323-873A-2

Query Match          30.5%; Score 717; DB 4; Length 339;
Best Local Similarity 54.9%; Pred. No. 8.5e-67;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLRGPVVAISLATEFFPYSFVADYIHPYANQSDPFYKPIELVNTKPLPVATTL 267
DB 67 LFPQHLPIKIAIATASLTFTYTLREVIHPLATSHQYFKPIIVINKVLPVNSITLL 126
QY 268 SLVYLAGLAAAYQLYGTYKRRPFWLETWLQCRKQGLISFFPNAVVAISLCIPMR 327
DB 127 ALVYLPVYIAIVQLHNGTKYKPPHLDKWLTRKQFGLISFFFAVIAIYSLSPMR 186
QY 328 SRVYFLNMAVOOYHANIENSMNEEVRIMYISFGIMSLGLSLATVTSVSNALN 387
DB 187 STRYKLMWAIQVQONKEDAMIEDYKMEIYSLGVGLATLALAVSISVSDSLT 246
QY 388 WREFSFIQSTLGYVALLSTFHVLIYGWKRAFEERYRFTPPNFVLAIVLPSIVIL 444
DB 247 WREFYIQSKIGIVISLGLTIALIFAMNKWIDIKQVWYTPPFMAVLPVIAI 303

RESULT 4
US-09-685-166A-879
Sequence 879, Application US/09685166A
Patent No. 6630305
GENERAL INFORMATION:
APPLICANT: Xu, Jianshun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqun
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Ranger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darriek
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C21
CURRENT APPLICATION NUMBER: US/09/685,166A
CURRENT FILING DATE: 2000-10-10
NUMBER OF SEQ ID NOS: 898
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 879
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
US-09-685-166A-879

Query Match          30.5%; Score 717; DB 4; Length 339;
Best Local Similarity 54.9%; Pred. No. 8.5e-67;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLRGPVVAISLATEFFPYSFVADYIHPYANQSDPFYKPIELVNTKPLPVATTL 267

```

```

Db      67 LEPQWMLPIKIAIILSLFTYLLREVIHPLATSHQYFYKIPILVINKVLPMVSIITLL 126
Qy      266 SLVYIAGLLAAAYOCYYGFKYRPPMLETMLQCKKOLLGSLFFPAMWYVSLCLPMR 327
      127 ALVYLPQVYAAVQVLANGRKFKFPHMDKMLTKQKGLSLFFFAVLAHAYLSYEMR 186
Qy      328 SRRYLFNAAVQVQVHANINENSWNEEYVRIEMYSISFGIMSLGLSLAVTSIPSVNALN 387
      187 SRRYKLIMAAVQVQVQNKEDAWIEHDVPMREIYVSLGIVGAIILALAVTSIPSDSLT 246
Qy      388 WEEPSIOSTLGIVALLISTFHVLLYGMKRAFEERYFYTPPENVALVPSIVIL 444
      247 WEEFHTQSKLGIVSLGLSTHALLFANKMIDIQFVWYTPPTFMIVFLPIVILI 303
Db

RESULT 5
US-09-655-270A-15
; Sequence 15, Application US/09655270A
; Patent No. 6329151
; GENERAL INFORMATION:
; APPLICANT: ROUVIERE, PIERRE E.
; TITLE OF INVENTION: High Density Sampling of Differentially Expressed Prokaryotic mR
; FILE REFERENCE: BC1011 US NA
; CURRENT APPLICATION NUMBER: US/03/655,270A
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/120,702
; PRIOR FILING DATE: 1999-February-19
; PRIOR APPLICATION NUMBER: 60/152,542
; PRIOR FILING DATE: 1999-September-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; TYPE: PRF
; LENGTH: 227
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-655-270A-15

Query Match
Best Local Similarity 28.0%; Score 144.5; DB 4; Length 227;
Matches 60; Conservative 33; Mismatches 88; Indels 33; Gaps 9;

Qy      25 IKDARKYTVGVIGSGDFAKSLTIRLCGYHYVIGSNPKFASFPFH-----VVD 75
      1 MSSSKIAVVG--GTGPGKGLAVFPAAGWPVIGSRSABAEALAEVRRAAGDAVVS 58
Db      76 VTHHEDALTKTNIIIPVAT---HREHYTSLMDLRHLVKGKILI-----DVNNMRIN 123
      59 AADNNSAADAQPIILLVVPYDGHRELVS---ELAFIPAGKLVVSCVNPGLGDSGAYGLD 115
Qy      124 QYPESSNAEYIASLFPDLSIVKGFNVVSA---WALQLGPKDASROYTICSNNOIARQOVIE 180
      116 VEEGSAAEQLRDVLPQATVVAAFHHLISAVNLMEHE--GP--LPEDVLYCGDDRSAGDEVAR 172
Db      181 LARQCNFIP-IDLSLSAREIENLPLRLFTLWR 213
Qy      173 LAVAITGRPGIDGALRVARQLEPLTAVLINVR 206
Db

RESULT 6
US-09-651-941-17
; Sequence 17, Application US/09651941
; Patent No. 6355470
; GENERAL INFORMATION:
; APPLICANT: ROUVIER, PIERRE E
; APPLICANT: WALTERS, DANA M
; APPLICANT: RAINIER, RUSS
; TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
; FILE REFERENCE: BC1022 US NA
; CURRENT APPLICATION NUMBER: US/09/651,941
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,545
; PRIOR FILING DATE: 1999-10-03
; NUMBER OF SEQ ID NOS: 28
US-09-651-941-17

```

```

; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 227
; TYPE: PRF
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-651-941-17

Query Match
Best Local Similarity 28.0%; Score 144.5; DB 4; Length 227;
Matches 60; Conservative 33; Mismatches 88; Indels 33; Gaps 9;

Qy      25 IKDARKYTVGVIGSGDFAKSLTIRLCGYHYVIGSNPKFASFPFH-----VVD 75
      1 MSSSKIAVVG--GTGPGKGLAVFPAAGWPVIGSRSABAEALAEVRRAAGDAVVS 58
Db      76 VTHHEDALTKTNIIIPVAT---HREHYTSLMDLRHLVKGKILI-----DVNNMRIN 123
      59 AADNNSAADAQPIILLVVPYDGHRELVS---ELAFIPAGKLVVSCVNPGLGDSGAYGLD 115
Qy      124 QYPESSNAEYIASLFPDLSIVKGFNVVSA---WALQLGPKDASROYTICSNNOIARQOVIE 180
      116 VEEGSAAEQLRDVLPQATVVAAFHHLISAVNLMEHE--GP--LPEDVLYCGDDRSAGDEVAR 172
Db      181 LARQCNFIP-IDLSLSAREIENLPLRLFTLWR 213
Qy      173 LAVAITGRPGIDGALRVARQLEPLTAVLINVR 206
Db

RESULT 7
US-09-955-597-17
; Sequence 17, Application US/09955597
; Patent No. 6461856
; GENERAL INFORMATION:
; APPLICANT: ROUVIER, PIERRE E
; APPLICANT: WALTERS, DANA M
; APPLICANT: RAINIER, RUSS
; TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
; FILE REFERENCE: BC1023 US NA
; CURRENT APPLICATION NUMBER: US/09/955,597
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/152,545
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; TYPE: PRF
; LENGTH: 227
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-955-597-17

Query Match
Best Local Similarity 28.0%; Score 144.5; DB 4; Length 227;
Matches 60; Conservative 33; Mismatches 88; Indels 33; Gaps 9;

Qy      25 IKDARKYTVGVIGSGDFAKSLTIRLCGYHYVIGSNPKFASFPFH-----VVD 75
      1 MSSSKIAVVG--GTGPGKGLAVFPAAGWPVIGSRSABAEALAEVRRAAGDAVVS 58
Db      76 VTHHEDALTKTNIIIPVAT---HREHYTSLMDLRHLVKGKILI-----DVNNMRIN 123
      59 AADNNSAADAQPIILLVVPYDGHRELVS---ELAFIPAGKLVVSCVNPGLGDSGAYGLD 115
Qy      124 QYPESSNAEYIASLFPDLSIVKGFNVVSA---WALQLGPKDASROYTICSNNOIARQOVIE 180
      116 VEEGSAAEQLRDVLPQATVVAAFHHLISAVNLMEHE--GP--LPEDVLYCGDDRSAGDEVAR 172
Db      181 LARQCNFIP-IDLSLSAREIENLPLRLFTLWR 213
Qy      173 LAVAITGRPGIDGALRVARQLEPLTAVLINVR 206
Db

RESULT 8
US-08-487-886-2

```

```

Sequence 2, Application US/08487886
Patent No. 5744448
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweickhardt, Rene Lynn
APPLICANT: Cheng, Shirley Yui Yen
APPLICANT: Nugent, No. 5744448een Patrice
TITLE OF INVENTION: Human Follicle Stimulating
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Stephan P. Williams,
ADDRESSEE: Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massll via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,886
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448

```

```

IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-487-886-2

Query Match 4.6%; Score 107.5; DB 1; Length 695;
Best local Similarity 17.0%; Pred. No. 0.028;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

QY 79 HEDALTKNTIIVVAHREHYTSLMDRLVVG-----KILIVSNMR 121
DB 98 HETRIKANKNL-LVYNPEAFQNLVNYLLSNTGIRKLPVYKHISLQVLDIQNIN 156
QY 122 INQYPSNAEYLASFPDSLIYKGFNVVSANALQGPDSARQVYICSNNOIARQVIEL 181
DB 157 IH-----TIERNFVGLSFESVILWL-----NNGGQIEHNCA----- 189
QY 182 ARGLNRPIDGLSSAREIEULPIRLFTWRGPVVAISLAFPIYSPVVDYIHRYAR 241
DB 190 ---FNGTQDLDELNDNNNLELPVDPVFGASGPVLDISRTIRHSLPSYLENWKLR 246
QY 242 NQSDPEYKIPRIYVKTLLPIVATITLSLVY-----PWTETWQCRKQL 305
DB 247 RSTYNNKILP-----TEKLVALMEASLTYPSCAFANMRROISELHPICNSILROE 300
QY 272 -----LAGLLAAAYQLYGTGYRRF-----PWTETWQCRKQL 305
DB 301 VDMTQTRQGRSLAEEDNESSYRSGFDMTYTEFYDCLNEVYDVYTCSPKPAFPCEDIM 360
QY 306 G---LISFFPAMVVA-----YSICLPFRSERYLFLNMAYQQVHANIN 347
DB 361 GYNILRLVLIWFTSILATIGNIIVLITTSQYKLTVP-----RELMDNLAFAD----- 408
QY 348 SNEEESVMIEMVYISGMSIGLLSLAVTSIPSVSNLNR-----PFIQSTL 398
DB 409 -----LCIGYLLIABVDHITSQHNVAIDMQTAGDAGFTVFASEL 455
QY 399 G--VYALISTFR-----VLIGMKCAFEERYRYFTPNFV-LA 435
DB 456 SVYTLTAITLERWHTITHAMQDCKVQLRHAASVWGMIFAFAALPIFGISSYKVS 515
QY 436 LVLPSTVLDLQ 449
DB 516 ICLPMDIDSPISQL 529

```

RESULT 9
US-08-482-855-2
Sequence 2, Application US/08482855
Patent No. 6121016
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schwaiblmair, Rene Lynn
APPLICANT: Cheng, Shirley Yui Yen
APPLICANT: Nugent, No. 6121016een Patricia
TITLE OF INVENTION: Human Follicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESS: Area-Serono, Inc.,
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,855
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length

FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative carboxy-terminal intracellular
LOCATION: 614 to 678
US-08-482-855-2
Query Match
Best Local Similarity 4.6%; Score 107.5; DB 3; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.028;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;
QY 79 HEDATKNIIFVIAHREHYTSLMDLRLHYG-----KILDVSNMR 121
DB 98 HEIRIEKANNL-LVINPFAFQNLNLOYLASNGIKHLPDVKIKSLQKVLIDIDQNN 156
QY 122 INQYPSNAEYLALFPDSLLVKGPNVSAWALQGPXASROYVICSNNIQARQVIEL 181
DB 157 IH-----TIENSFVGLSFESEVILWL-----NKNGICEIHNC----- 189
QY 182 AQLNFIPIDIGSSASAEINLFLRLPTLMRGPVVAISLATFFFLYFVRDVIHPYAR 241
DB 190 ---FNGTDLDELANSDDNNLEBELPNDVPHGASGVILIDIRTHSLPSYGLNKKLRA 246
QY 242 NOQSDFYKIPLEIVNKL-PIVAITLSLVY----- 271
DB 247 RSTYVKKLP-----TLEKVALMEASLTYPHSCAFANMRQISELHPICNKSLRQE 300
QY 272 -----TAGLAAAYQLYGTGYRRF-----PPWLETWLQCRKOL 305
DB 301 VDMYQTRGQRSLSAEDNESYSRGFDMYTEPFDYD-CNEVDVTCSPKFDANPCEDIM 360
QY 306 G-----LSFFPAWYHVA-----YSLCLPMRSEVYLPLNMAVQOVHANLEN 347
DB 361 GNIIRVLIWFTSIALATGNIIVILITQYKLTVP-----RFLMCLAFAD----- 408
QY 348 SWNEBEVVRIMYISFGIMSLGLSLAVTISPSVNALMRE-----FSFIQSTL 398
DB 409 -----LCIGIYLLIASVDIHFKSYHNAIIMQAGACDAGFTVASEL 455
QY 399 G-----YVALLISTFH-----TLIYMKRABEEYRYRTPENY-IA 435
DB 456 SYTTITATILERWHTITTAQMDCKVQLRAASVWVGWIFAPRAALFPIGISITWKS 515
QY 436 LVLPSIVLIDLQL 449

Db 516 ICLPMDIDPSLSQL 529

RESULT 10
US-08-474-986-2
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
Schweickhardt, Rene Lynn
Cheng, Shirley Yui Yen
Nugent, No. 6372711een Patrice
TITLE OF INVENTION: Human Follicle Stimulating
Hormone Receptor

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
Ares-Serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Masell via Kermit to IBM MS-DOS

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474, 986
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670, 085
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
LOCATION: 6.4 to 678

SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-474-986-2

Query Match 4.6%; Score 107.5; DB 4; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.028;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

Db 79 HEDALTKTNIIFVAIHREHYTSLMDRLHLLVG-----KILIDVSNMR 121
122 INGVESNNAVYLAFLPDSLIYKGFVVSAMALQGPXKASQVYICNNIQAQGVTEL 181
157 IH-----TIERNSEVCSIESVILML-----NKGIOEINCA----- 189
182 ARQNLPIPIDLGSLASAREIENPLRLFTLMRGPVVAISLATPFVYSFVADVIHPYAR 241
190 ---FNGQDELNLSDNNLELPNDVFHAGSGPVLDIRTRISLPSYGLENKKLRA 246
242 NQSDFKYKPIEVNKTLLPIVATILLSIVY----- 271
247 RSTYNKKL-----TLEKVALMEASLTYPSHCAPANMRROISELHPICNSILRGE 300
272 -----LAGLAAAYQLYGYTKYRR-----PWELEWLQCRKQL 305
301 VDYWTQRCORSSIAEDNESSYSGRPFDMYTEVDYDLCNEVVDYTCSPDPDFNCEIDIM 360
306 G-----LISFFPAVAVHA-----YSLCLPMRSEBRYLFLNMAVOOVHANIE 347
361 GYNLRKLVINFIISLAIATGNIIYLVITTSQYKLTVP-----RFLMCKNLAFAD----- 408

Db 348 SWNEEYWRIMYISFGIMSLGLSLAVTSIPSVSNALNWR-----FSFIQSTL 398
409 -----LCITGIYLLILNASVDIHTKSOYHNHAIIMQAGCDAAAGFTVASEL 455

Db 399 G---YVALLISTH-----VLIYGRKAPFEERYRPTPEV-IA 435
456 SVYTLTATLTERWHTTHAMQDCKVQLRRAASVYMWGWIIFAPPAALFPIFGISSYKVS 515

Db 436 LVLPSTVILDLQL 449
516 ICLPMDIDPSLSQL 529

RESULT 11
US-09-323-873A-20
Sequence 20, Application US/09323873A
Patent No. 6329503
GENERAL INFORMATION:
APPLICANT: Daniel E. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Kahan Leong
APPLICANT: Arthur B. Raitano
APPLICANT: Douglas C. Saffran
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
FILE REFERENCE: 129.16SU2
CURRENT APPLICATION NUMBER: US/09/323, 873A
CURRENT FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087, 520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091, 183
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 20
LENGTH: 34
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: STEAP-1 PEPTIDE

US-09-323-873A-20

Query Match 4.6%; Score 107; DB 4; Length 34;
Best Local Similarity 58.8%; Pred. No. 0.0003;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Db 326 RRSERYLFLNAYQOVHANIESSWNEEYWRIM 359
1 RRSRYKKLINMAYQOVQNKEDAVIEHDVWRMET 34

RESULT 12
US-09-328-352-8165
Sequence 8165, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328, 352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 8165
LENGTH: 940
TYPE: PRT
ORGANISM: Acinetobacter baumannii

US-09-328-352-8165

Query Match 4.4%; Score 102.5; DB 4; Length 940;
Best Local Similarity 19.5%; Pred. No. 0.15;
Matches 71; Conservative 74; Mismatches 146; Indels 73; Gaps 17;

Fri Apr 16 16:26:35 2004

us-09-455-486-6.rai

Page 8

```

Qy      132 YLASJEPDLSIVKGFVNVSAAMALQIGPXDASQVYIGSNNIQA-----ROOVLEIRCOLN 186
Db      298 YLSVJVPDLSI-SKTLANVAIVAVLCGTLSA---LCYISLSLSGRRQALDIRQA 352
Qy      187 FIFIDL-GSISAREIENTPLRLFTIMRGPVVAISLA-----TFPFLXSPVDYIHP 238
Db      353 FRPLWIGLSIAALGEVAHPRLIAGIGENTSICISLTANASALFTALFWRRFRPPAH- 411
Qy      239 YARNOQ-----SDPKYKPI-----EIVNK 257
Db      412 LIRNPRLRLKRRSLHDVLQVGSIMFPVLVYVGLSPATFVSAQSSALRLVLCA 471
Qy      258 TLPIVATLISLVYLAAGLAAAYQLYGKYYRRFPFMTLQCRKQUGLSPFPAWY 317
Db      472 VLAVVAVMTVGLL-----RRSSVWGQPPRSAPYIQ-----LOSFOYTLIHI 515
Qy      318 AYSCLFMRRSERYLFPLNVAVQOVHAININSNNEEYVRIEY-YISPGIM----- 366
Db      516 FVVVL-----FPIEVALRWGSGSLIR-YAESEGRQISMKVVSFGFTLLVAMIL 564
Qy      367 -----SLGL-----LSJLAIVSISVSNALNWRSPSIOSTLGVYLLIS 406
Db      565 TDPAIQSLSGIGGKSRBNTRALMTMLIRNVL-----FATIVIALIYA 608

```

Search completed: April 16, 2004, 12:41:47
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:32:36 ; Search time 60 Seconds
(without alignments)
2137.943 Million cell updates/sec

Title: US-09-455-486-6

Sequence: 1 MBSISWMSGPKSLSETCLPN.....ALVPSIVLLDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20048:*
8: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2351	100.0	454	AAE02781	Human six
2	2351	100.0	454	AAU10188	Human ORF
3	2351	100.0	454	ABG61933	Prostate
4	2351	100.0	454	AAU76538	Tumour-as
5	2351	100.0	454	AAU80190	Human PUM
6	2351	100.0	454	AAU08893	Tumour-as
7	2348	99.9	454	AAU65001	Human pro
8	2294	97.6	490	AAE28951	Human STE
9	2294	97.6	490	AAU63112	Human tra
10	2290	97.4	490	AAU10187	Human six
11	2036	86.6	419	AAU10189	Human ORF
12	1972.5	83.9	576	ABG12306	Novel hum
13	1972.5	83.9	1273	ABG00113	Novel hum
14	1738	73.9	450	AAE02841	Human STE
15	1364.5	58.0	450	ABP62883	Human pol
16	1316	56.0	488	AAAB49483	Rat p-Hyd
17	1316	56.0	526	AAAB83365	Murine Tu
18	1272	54.1	488	AAAB83365	Human trn
19	1272	54.1	488	AAAB83366	Human trn
20	1269	54.0	488	AAAB93224	Human pro
21	1269	54.0	488	AAU04564	Human G-P
22	1269	54.0	488	AAU10220	Human Six
23	1269	54.0	488	ABU60864	Human G-P
24	1245.5	53.0	487	AAAB49481	Human p-H
25	1088	46.3	459	AAAB74715	Human mem

26	1087	46.2	459	AAE02782	Human six
27	1087	46.2	459	ABP64820	Human pro
28	1085	46.2	459	AAU10190	Human six
29	1082	46.0	458	AAE02636	Human STE
30	1074.5	45.7	456	AAAB49482	Human p-H
31	901	38.3	173	AAAB8195	Human PUM
32	844	35.9	179	AAAB18006	peptide #
33	844	35.9	179	AAAB37041	peptide #
34	844	35.9	179	AAAB30517	peptide #
35	844	35.9	179	AAAB31810	peptide #
36	844	35.9	179	ABAB22356	protein #
37	844	35.9	179	AAAW0183	Human bon
38	844	35.9	179	AAAB7768	Human bra
39	844	35.9	179	ABG51883	Human liv
40	844	35.9	179	AAAB05646	peptide #
41	844	35.9	179	ABG39817	Human pep
42	768	32.7	237	AAU04565	Human G-P
43	768	32.7	237	ABU60865	Human G-P
44	736	31.3	141	AAAY52589	Human pro
45	717	30.5	267	ABU60886	Human G P

ALIGNMENTS

RESULT 1	
AAE02781	AAE02781 standard; protein; 454 AA.
XX	XX
AC	AAE02781;
XX	XX
DT	06-AUG-2001 (first entry)
XX	XX
DE	Human six transmembrane epithelial antigen of prostate (STEAP)-2 protein.
XX	XX
KW	Human; cytosolic; antiproliferative; vaccine; gene therapy.
KW	six transmembrane epithelial antigen of the prostate-2; STEAP-2;
KW	chromosome 7q21; cancer; prostate; colon; bladder; lung; ovarian;
KW	pancreatic.
XX	XX
OS	Homo sapiens.
XX	XX
EH	Key
FT	Region
FT	/label= HLA-A2_binding_peptide #5
FT	153.165
FT	/label= Immunogenic_peptide #1
FT	207.228
FT	/label= Transmembrane_domain #1
FT	227.235
FT	/label= HLA-A2_binding_peptide #1
FT	255.277
FT	/label= Transmembrane_domain #2
FT	304.325
FT	/label= Transmembrane_domain #3
FT	306.314
FT	/label= HLA-A2_binding_peptide #4
FT	307.315
FT	/label= HLA-A2_binding_peptide #3
FT	345.358
FT	/label= Immunogenic_peptide #2
FT	359.381
FT	/label= Transmembrane_domain #4
FT	393.415
FT	/label= Transmembrane_domain #5
FT	402.410
FT	/label= HLA-A2_binding_peptide #2
FT	428.450
FT	/label= Transmembrane_domain #6
XX	XX
PN	W0200140276-A2.
XX	XX
PD	07-JUN-2001.

QY 121 RINQYPSNNAEYLAFLPDSLIYKGFNVSAWALQGPDKASROYICSNNIQARQVIE 180
 CC associated polynucleotides (designated PC genes) that selectively
 CC hybridise to a sequence that is at least 80% identical to them. The
 CC prostate cancer-associated polynucleotide sequences are differentially
 CC expressed in prostate tumour tissue or in prostate cancer and are derived
 CC from the tissues of various organisms such as humans or other mammals
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
 CC associated genes are useful for diagnosing or treating prostate cancer,
 CC as well as for identifying modulators of prostate cancer or agents that
 CC inhibit prostate cancer. The nucleic acid sequences are particularly
 CC useful in gene therapy, as a vaccine or in antisense applications.
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins

QY 121 RINQYPSNNAEYLAFLPDSLIYKGFNVSAWALQGPDKASROYICSNNIQARQVIE 180
 Db 121 RINQYPSNNAEYLAFLPDSLIYKGFNVSAWALQGPDKASROYICSNNIQARQVIE 180
 QY 181 LARQINFIPIIDLGSLSSAREIENPLRLFTLMRGPVVAISLAFPLYSFVRDIHPYA 240
 Db 181 LARQINFIPIIDLGSLSSAREIENPLRLFTLMRGPVVAISLAFPLYSFVRDIHPYA 240
 QY 241 RNOQSDFYKIPFIEIVNKTLPPIVAITLLSLVYLAGLAAAYQLYGKTRRPPMLETMO 300
 Db 241 RNOQSDFYKIPFIEIVNKTLPPIVAITLLSLVYLAGLAAAYQLYGKTRRPPMLETMO 300
 QY 301 CRKQGLISLFFPAMVAVASLCLPMRSEERYLFNNAYQOVHANIENSWNEEVRILEMY 360
 Db 301 CRKQGLISLFFPAMVAVASLCLPMRSEERYLFNNAYQOVHANIENSWNEEVRILEMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFIOSTLGYVALLISTFHVLIYGMKRAFE 420
 Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFIOSTLGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EBYRFFYTPPNFVALVLPISIVILDLQCRYPD 454
 Db 421 EBYRFFYTPPNFVALVLPISIVILDLQCRYPD 454

RESULT 3
 ABG61933
 ID ABG61933 standard; protein; 454 AA.
 XX ABG61933;
 AC 15-AUG-2002 (first entry)
 XX 15-AUG-2002 (first entry)
 DE Prostate cancer-associated protein #134.
 XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
 KM Mammalia.
 OS WO200230266-A2.
 XX 18-APR-2002.
 PD 12-OCT-2001; 2001WO-US032045.
 XX 12-OCT-2001; 2001WO-US032045.
 PF 13-OCT-2000; 2000US-00687576.
 XX 13-OCT-2000; 2000US-00687576.
 PR 08-DEC-2000; 2000US-00733288.
 XX 08-DEC-2000; 2000US-00733288.
 PR 24-JAN-2001; 2001US-0263957F.
 XX 24-JAN-2001; 2001US-0263957F.
 PR 16-MAR-2001; 2001US-0276791B.
 XX 16-MAR-2001; 2001US-0276791B.
 PR 06-APR-2001; 2001US-0281922P.
 XX 06-APR-2001; 2001US-0281922P.
 PR 24-APR-2001; 2001US-0286214P.
 XX 24-APR-2001; 2001US-0286214P.
 PR 30-APR-2001; 2001US-00847046.
 XX 30-APR-2001; 2001US-00847046.
 PR 04-MAY-2001; 2001US-0286589P.
 XX 04-MAY-2001; 2001US-0286589P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
 XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
 PI WPI; 2002-471335/50.
 XX N-PSDB; ABK92252.
 DR N-PSDB; ABK92252.
 XX Detecting a prostate cancer-associated transcript in a cell in a patient,
 PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
 PT by determining if prostate cancer-associated genes are expressed in a
 PT prostate tissue.
 XX Claim 27; Page 416; 436pp; English.
 PS The present invention relates to methods of detecting a prostate cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with prostate cancer-

CC associated polynucleotides (designated PC genes) that selectively
 CC hybridise to a sequence that is at least 80% identical to them. The
 CC prostate cancer-associated polynucleotide sequences are differentially
 CC expressed in prostate tumour tissue or in prostate cancer and are derived
 CC from the tissues of various organisms such as humans or other mammals
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
 CC associated genes are useful for diagnosing or treating prostate cancer,
 CC as well as for identifying modulators of prostate cancer or agents that
 CC inhibit prostate cancer. The nucleic acid sequences are particularly
 CC useful in gene therapy, as a vaccine or in antisense applications.
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins

QY 121 RINQYPSNNAEYLAFLPDSLIYKGFNVSAWALQGPDKASROYICSNNIQARQVIE 180
 Db 121 RINQYPSNNAEYLAFLPDSLIYKGFNVSAWALQGPDKASROYICSNNIQARQVIE 180
 QY 181 LARQINFIPIIDLGSLSSAREIENPLRLFTLMRGPVVAISLAFPLYSFVRDIHPYA 240
 Db 181 LARQINFIPIIDLGSLSSAREIENPLRLFTLMRGPVVAISLAFPLYSFVRDIHPYA 240
 QY 241 RNOQSDFYKIPFIEIVNKTLPPIVAITLLSLVYLAGLAAAYQLYGKTRRPPMLETMO 300
 Db 241 RNOQSDFYKIPFIEIVNKTLPPIVAITLLSLVYLAGLAAAYQLYGKTRRPPMLETMO 300
 QY 301 CRKQGLISLFFPAMVAVASLCLPMRSEERYLFNNAYQOVHANIENSWNEEVRILEMY 360
 Db 301 CRKQGLISLFFPAMVAVASLCLPMRSEERYLFNNAYQOVHANIENSWNEEVRILEMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFIOSTLGYVALLISTFHVLIYGMKRAFE 420
 Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFIOSTLGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EBYRFFYTPPNFVALVLPISIVILDLQCRYPD 454
 Db 421 EBYRFFYTPPNFVALVLPISIVILDLQCRYPD 454

RESULT 4
 AAU76538
 ID AAU76538 standard; protein; 454 AA.
 XX AAU76538;
 AC 05-JUN-2002 (first entry)
 XX 05-JUN-2002 (first entry)
 DE Tumour-associated antigenic target protein, TAT138.
 XX Tumour-associated antigenic target protein, TAT138.
 KM TAT138; Tumour-associated Antigenic Target; tumour; breast cancer;
 KM colorectal cancer; lung cancer; ovarian cancer;
 KM central nervous system cancer; liver cancer; bladder cancer; melanoma;
 KM pancreatic cancer; leukaemia; gene therapy.
 XX Homo sapiens.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH Modified-site 8..13
 FT Modified-site /note="N-myristoylation site"
 FT Modified-site 24..29

PI Hillan KJ, Watanabe CK, Wood WI, Yansura DG, Zhang Z;
XX WPI; 2002-383270/41.
DR N-PSDB; ABRK0391.
XX
PT New polypeptide termed protein upregulated in metastatic prostate cancer
PT and encoding polynucleotides, useful for identifying polypeptide
PT antagonists for treating prostate cancer.

PS Claim 23; Fig 2; 137pp; English.

The invention relates to an isolated human protein upregulated in metastatic prostate cancer (PUMPCn). PRO32303 polypeptide, a sequence 80% identical to PRO32303 and the sequence as encoded by cDNA insert of the vector deposited as ATCC Deposit No. PTA-2513 (DNA185171- 2994) on 26/9/2000. Also included are the polynucleotide encoding the protein (or a DNA sequence 80% identical to the polynucleotide) and one that hybridises to complement of the polynucleotide), a vector comprising the polynucleotide, a polynucleotide deposited with ATCC under accession number PTA-2513 (DNA185171-2994), a host cell comprising the vector, preparation of PRO32303, a chimeric molecule comprising PRO32303 fused to a heterologous amino acid sequence, an anti-PRO32303 antibody, an agonist/antagonist of PRO32303, and diagnosing the presence of prostate cancer in a mammal by: (a) contacting a microarray diagnostic with a DNA185171-2994 probe, detecting and quantifying hybridisation of DNA185171-2994 probe in prostate cancer tissue compared with normal tissue and determining if DNA185171-2994 is overexpressed; or (b) contacting a tissue of the mammal with an anti-PRO32303 antibody and detecting the binding of the antibody to a component of the tissue, where binding is indicative of the presence of prostate cancer in the mammal. The antibody is useful for treating prostate cancer in mammal which is androgen independent prostate cancer, that has metastasised to another portion of the body, where the antibody is not conjugated with a cytotoxic agent and the method further comprises administering a chemotherapeutic agent to the mammal and for diagnosing the presence of prostate cancer in a mammal. The PRO32303 polynucleotide is useful for generating either transgenic animals or knock out animals which in turn are useful in the development and screening of therapeutically useful reagents or for use in gene therapy and for chromosome identification. PRO32303 and polynucleotide may also be used for tissue typing and PRO32303 may also be employed as a therapeutic agents and for screening compounds to identify those that mimic the PRO32303 polypeptides (agonists) or prevent the effect of the PRO32303 polypeptide (antagonists). PRO32303 is further useful for the affinity purification of PRO32303 from recombinant cell culture or natural sources. The present sequence represents PRO32303

SQ Sequence 454 AA;

Query Match	100.0%	Score 2351	DB 5	length 454
Best Local Similarity	100.0%	Pred. No. 5e-245		
Matches 454; Conservative	0	Mismatches	0	Indels 0; Gaps 0

QY	1	MSISIMWGGSPKSLSEKCLNGINGINGIDAKRVGVIGSDCPKXSLTITLRCGYNVIGS	60
Db	1	MSISIMWGGSPKSLSEKCLNGINGINGIDAKRVGVIGSDCPKXSLTITLRCGYNVIGS	60
QY	61	RNPKFASPEPPHYVDYTHEDDALTKNII FVAIHREHYTSLMDLRHLVGLKILIDVSNM	120
Db	61	RNPKFASPEPPHYVDYTHEDDALTKNII FVAIHREHYTSLMDLRHLVGLKILIDVSNM	120
QY	121	RINQYEPESNAEYIASLPDDEL IYKGRNVSAMALQCPDASQVYICGNNIQARQVIE	180
Db	121	RINQYEPESNAEYIASLPDDEL IYKGRNVSAMALQCPDASQVYICGNNIQARQVIE	180
QY	181	LAQNLNFIPIDLGSLSGAREIENLPRLFLTMGCPVYVAISLATPFFIYSFVRDVTHHPA	240
Db	181	LAQNLNFIPIDLGSLSGAREIENLPRLFLTMGCPVYVAISLATPFFIYSFVRDVTHHPA	240
QY	241	RNQQSDFYKPIEIVNNTLPIVAITLTLSLVYLAGLLAAAYQLYYGKTKRRFPFWLETWQ	300
Db	241	RNQQSDFYKPIEIVNNTLPIVAITLTLSLVYLAGLLAAAYQLYYGKTKRRFPFWLETWQ	300

QY	301	CRQQLGLSFFPAMVAVASLCIPMR3SERLPLNNAYOOVHANIEENSNKEEYWRLEMY	360
Db	301	CRQQLGLSFFPAMVAVASLCIPMR3SERLPLNNAYOOVHANIEENSNKEEYWRLEMY	360
QY	361	ISRGIMSLGLSLTAVTSIPSVSNALNMRREPSFQSTLTGVALLISTFHYLIYGMKRAFE	420
Db	361	ISRGIMSLGLSLTAVTSIPSVSNALNMRREPSFQSTLTGVALLISTFHYLIYGMKRAFE	420
QY	421	EEYRREYTPPNFVALVLPISIVILDLQLCRYPD	454
Db	421	EEYRREYTPPNFVALVLPISIVILDLQLCRYPD	454

RESULT 6

ID ABU08893 standard; protein; 454 AA.

DT 09-OCT-2003 (first entry)

DE Tumour-associated antigenic target 138 (TAT138).

XX Human; tumour-associated antigenic target polypeptide; TAT; tumour;
 KM gene therapy; cytostatic; tissue typing; prostate tumour; cancer.

OS Homo sapiens

PN US2003060612-A1.

PD 27-MAR-2003

XX
PF 22-JUN-2001: 2001US-00888257.

XX		
PR	28-OCT-1997;	97US-0063540P
PR	17-UTN-1998;	98US-0086563P
PR	10-SEP-1998;	98US-0099792P
PR	08-OCT-1998;	98US-0103618P
PR	02-JUN-1999;	99MC-US0121252
PR	01-SEP-1999;	99MC-US0201111
PR	18-FEB-2000;	2000MC-US0043432
PR	02-MAR-2000;	2000MC-US005841
PR	30-MAR-2000;	2000MC-US008439
PR	24-SEP-2000;	2000MC-US0233528
PR	26-SEP-2000;	2000US-0235453P
PR	01-DEC-2000;	2000MC-US0326278
PR	28-FEB-2001;	2001MC-US0066520
PR	01-MAR-2001;	2001MC-US0066666

XX
PA (GETH) GENENTECH INC.

xx Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V,
 PI Wood WI, Wu TD, Zhang Z,
 PI

XX	New nucleic acid, useful for the manufacture of a medicament for
PT	diagnosing or treating tumor in a mammal.
PT	

PS Claim 12; Fig 10; 71pp; English.

XX The invention discloses human nucleic acids encoding tumour-associated antigenic target (TAT) polypeptides, with or without their associated signal peptide. Also disclosed is an antibody that specifically binds to the TAT polypeptides, a method for detecting the presence of a tumour in a mammal and a method for killing a cancer cell expressing the TAT polypeptide. The nucleotide sequences are useful in preparing TAT polypeptides by recombinant techniques and in gene therapy (e.g. for replacement of defective gene). The TAT polypeptides are useful as therapeutic agents and for detecting the presence, prevention and/or treatment of a tumour, such as colon, breast or prostate tumour. The TAT polypeptides and nucleic acids may also be used diagnostically for tissue

CC Typing. The sequence presented is the TAT138 polypeptide of the invention
 XX SQ Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 7; Length 454;
 Best Local Similarity 100.0%; Pred. No. 5e-245;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWGPXKSLSETCLPENGINGIKDARKVTGVYIGSGDPKSLITILNCGHVYIGS 60
 DB 1 MESISMWGPXKSLSETCLPENGINGIKDARKVTGVYIGSGDPKSLITILNCGHVYIGS 60
 QY 61 RNPKEASEFPFHVVDVTHHEDALTKNIIFFVAIHRHRYTSLMDLRHLVKGILLIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHHEDALTKNIIFFVAIHRHRYTSLMDLRHLVKGILLIDVSNM 120
 QY 121 RINQYBESNAEYIASLPDGLIVKGNVVSAMALQGPDKASRQYICSNNIQARQVIE 180
 DB 121 RINQYBESNAEYIASLPDGLIVKGNVVSAMALQGPDKASRQYICSNNIQARQVIE 180
 QY 181 LARQNLFPIDGLISASAREIENLPLRLFTLMRGPVVVAISLATEFFLYSFVRDIHPYA 240
 DB 181 LARQNLFPIDGLISASAREIENLPLRLFTLMRGPVVVAISLATEFFLYSFVRDIHPYA 240
 QY 241 RNOQSDFFYKPIEIVNKTPIVAITLISLVYLAGLLAAAYQYGTGRFPPMLETWLQ 300
 DB 241 RNOQSDFFYKPIEIVNKTPIVAITLISLVYLAGLLAAAYQYGTGRFPPMLETWLQ 300
 QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYLFLNMAVQOVHANIENSNMEEVWRLEMY 360
 DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYLFLNMAVQOVHANIENSNMEEVWRLEMY 360
 QY 361 ISFGIMSLGLSLAVTISIPSVSNALNMRSEFIOSTLGYVALLISTFHLIYGMKKAPE 420
 DB 361 ISFGIMSLGLSLAVTISIPSVSNALNMRSEFIOSTLGYVALLISTFHLIYGMKKAPE 420
 QY 421 EBYRFFYPNPNFVLAIVPSIVILLDLQLCRYPD 454
 DB 421 EBYRFFYPNPNFVLAIVPSIVILLDLQLCRYPD 454

RESULT 7
 ADB65001
 ID ADB65001 standard; protein; 454 AA.
 XX AC ADB65001;
 XX DT 04-DEC-2003 (first entry)
 DE Human protein encoded by clone PROST20168600.
 XX KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
 KW cell regeneration; membrane protein; signal transduction-related protein;
 KW transcription-related protein; osteoporosis; neurological disease;
 KW cancer; tumor.
 XX OS Homo sapiens.
 XX PN EPI3064557a2.
 XX PD 07-MAY-2003.
 XX PF 28-MAR-2002; 2002EP-00007401.
 XX PR 05-NOV-2001; 2001JP-00379298.
 XX PR 25-JAN-2002; 2002US-00350978.
 XX PA (HELI-) HELIX RES INST.
 XX PA (RBAS-) RES ASSOC BIOTECHNOLOGY.
 XX PI Isegaki T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
 PI Yamamoto J, Isono Y, Hito Y, Otsuka K, Nagai K, Irie R, Tamechika I,
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-450961/43.
 DR N-PSDB; ADB63031.
 XX PT New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.

XX Claim 1, Page: 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesizing the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumors). The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a protein of the invention. Note: Some of the
 CC sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.

XX SQ Sequence 454 AA;

Query Match 99.9%; Score 2348; DB 7; Length 454;
 Best Local Similarity 99.8%; Pred. No. 1.1e-244;
 Matches 453; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWGPXKSLSETCLPENGINGIKDARKVTGVYIGSGDPKSLITILNCGHVYIGS 60
 DB 1 MESISMWGPXKSLSETCLPENGINGIKDARKVTGVYIGSGDPKSLITILNCGHVYIGS 60
 QY 61 RNPKEASEFPFHVVDVTHHEDALTKNIIFFVAIHRHRYTSLMDLRHLVKGILLIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHHEDALTKNIIFFVAIHRHRYTSLMDLRHLVKGILLIDVSNM 120
 QY 121 RINQYBESNAEYIASLPDGLIVKGNVVSAMALQGPDKASRQYICSNNIQARQVIE 180
 DB 121 RINQYBESNAEYIASLPDGLIVKGNVVSAMALQGPDKASRQYICSNNIQARQVIE 180
 QY 181 LARQNLFPIDGLISASAREIENLPLRLFTLMRGPVVVAISLATEFFLYSFVRDIHPYA 240
 DB 181 LARQNLFPIDGLISASAREIENLPLRLFTLMRGPVVVAISLATEFFLYSFVRDIHPYA 240
 QY 241 RNOQSDFFYKPIEIVNKTPIVAITLISLVYLAGLLAAAYQYGTGRFPPMLETWLQ 300
 DB 241 RNOQSDFFYKPIEIVNKTPIVAITLISLVYLAGLLAAAYQYGTGRFPPMLETWLQ 300
 QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYLFLNMAVQOVHANIENSNMEEVWRLEMY 360
 DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYLFLNMAVQOVHANIENSNMEEVWRLEMY 360
 QY 361 ISFGIMSLGLSLAVTISIPSVSNALNMRSEFIOSTLGYVALLISTFHLIYGMKKAPE 420
 DB 361 ISFGIMSLGLSLAVTISIPSVSNALNMRSEFIOSTLGYVALLISTFHLIYGMKKAPE 420
 QY 421 EBYRFFYPNPNFVLAIVPSIVILLDLQLCRYPD 454
 DB 421 EBYRFFYPNPNFVLAIVPSIVILLDLQLCRYPD 454

RESULT 8
AAE28951.
ID AAE28951 standard; protein; 490 AA.
XX
AC AAE28951;
XX
DT 27-JAN-2003 (first entry)
XX
DE Human STEAPRP protein.
XX
KW Human, STEAPRP; proliferative disorder; prostate hyperplasia; cancer;
KM six-transmembrane epithelial antigen of the prostate related protein;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200272596-A1.
XX
PD 19-SEP-2002.
XX
PF 07-MAR-2002; 2002MO-US007053.
XX
PR 09-MAR-2001; 2001US-00802520.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal PG, Faris M, Chen H, Ison CH;
XX
DR WPI; 2002-723311/78.
DR N-PSDB; AAD46429.
XX
PT New six-transmembrane epithelial antigen of the prostate-related protein
PT and nucleic acids, useful for diagnosing and treating prostate cell
PT proliferative disorders, particularly prostate hyperplasia and prostate
PT cancer.
XX
PS Claim 13; Page 54-55; 62pp; English.
XX
CC The invention relates to six-transmembrane epithelial antigen of the
CC prostate related protein (STEAPRP) and its corresponding nucleic acid.
CC STEAPRP DNA is used to diagnose and to treat prostate cell proliferative
CC disorders, particularly prostate hyperplasia and prostate cancer, and for
CC screening a library of molecules of compounds for specific binding
CC affinity. It is also used in gene therapy. STEAPRP is used to screen a
CC subject sample for antibodies, which specifically binds the protein, and
CC to prepare and purify a protein. The present sequence is human STEAPRP
CC protein.
XX
SQ Sequence 490 AA;
Query Match 97.6%; Score 2294; DB 5; Length 490;
Best Local Similarity 100.0%; Pred. No. 8.4e-239; Indels 0; Gaps 0;
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESISMMGSPKSLSETCLPENGINGIKDARKTVGVIGSGDFAKSLTIRIRCGYHVIGS 60
DB 1 MESISMMGSPKSLSETCLPENGINGIKDARKTVGVIGSGDFAKSLTIRIRCGYHVIGS 60
QY 61 RNPKPSEFFPRVVDVTHEDALTKNITFVAIHREHYSLMDLHLVGLKLLIDVSNM 120
DB 61 RNPKPSEFFPRVVDVTHEDALTKNITFVAIHREHYSLMDLHLVGLKLLIDVSNM 120
QY 121 RINQYESNAEYLASLFPDSLIYKGFNVSAVALQGPDAEROYICSNNOARQOYIE 180
DB 121 RINQYESNAEYLASLFPDSLIYKGFNVSAVALQGPDAEROYICSNNOARQOYIE 180
QY 181 LARQNLFIPIDGLSSARIEMLPRLTLWRGPPVVAISLATPFLYSFRDVIHYA 240
DB 181 LARQNLFIPIDGLSSARIEMLPRLTLWRGPPVVAISLATPFLYSFRDVIHYA 240
QY 241 RNOQSDFYKIPRIYVNTLPIVAITLLSLVYLAGLAAAYQOLYGTXYRPPMLETWLQ 300
DB 241 RNOQSDFYKIPRIYVNTLPIVAITLLSLVYLAGLAAAYQOLYGTXYRPPMLETWLQ 300

DB 241 RNOQSDFYKIPRIYVNTLPIVAITLLSLVYLAGLAAAYQOLYGTXYRPPMLETWLQ 300
QY 301 CRKQGLSFFPANVAVSYCLPWRREERYLFIMAYQOVHANIEENSWNEEYWRLEMY 360
DB 301 CRKQGLSFFPANVAVSYCLPWRREERYLFIMAYQOVHANIEENSWNEEYWRLEMY 360
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRERFSFIQSTGYVALLISTRHVLIYWKRAFE 420
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRERFSFIQSTGYVALLISTRHVLIYWKRAFE 420
QY 421 EBYRFRYTPPNFVALVPSIVIL 444
DB 421 EBYRFRYTPPNFVALVPSIVIL 444
RESULT 9
ABU63312
ID ABU63312 standard; protein; 490 AA.
XX
AC ABU63312;
XX
DT 18-SEP-2003 (first entry)
XX
DE Human transmembrane protein TMPL.
XX
KM Human; prostate tumour; lung tumour; TMPL; cancer.
XX
OS Homo sapiens.
XX
PN US2003064397-A1.
XX
PD 03-APR-2003.
XX
PF 24-JUL-2002; 2002US-00205267.
XX
PR 22-MAY-1998; 98US-00083521.
PR 16-SEP-1999; 99US-00397558.
PR 09-MAR-2001; 2001US-00802520.
PR 26-SEP-2001; 2001US-00963896.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Spancake XM, Rickert PK, Lal PG, Ison CH;
XX
DR WPI; 2003-540783/51.
DR N-PSDB; ACD27642.
XX
PT New transmembrane protein differentially expressed in prostate and lung
PT tumors, useful for diagnosing, staging, creating or monitoring
PT progression or treatment of cancer, preferably lung or prostate cancer.
XX
PS Claim 1; Fig 1; 49pp; English.
XX
CC The invention relates to a purified human protein comprising a
CC polypeptide having a fully defined transmembrane protein differentially
CC expressed in prostate and lung tumour (TMPL) sequence. The protein is
CC useful for detecting expression of the protein in a sample, where the
CC sample is lung or prostate tissue sample and the protein is
CC differentially expressed when compared with a standard and is diagnostic
CC of a lung or prostate cancer. The protein is useful to identify an
CC antibody that specifically binds the protein. The protein is useful for
CC screening several molecules and compounds to identify at least one
CC ligand. The protein is also useful for preparing and purifying a
CC polyclonal antibody and for preparing a monoclonal antibody using
CC standard animal immunisation/hybridoma methods. A TMPL specific antibody
CC is useful to detect expression of the protein in a lung or prostate
CC tissue sample and the method is useful as a diagnostic for lung or
CC prostate cancer. A TMPL specific antibody is also useful for
CC immunopurifying a protein, for treating a prostate cancer and for
CC delivering a therapeutic agent to a cancer, preferably prostate cancer.
CC The protein is useful for diagnosing, staging, treating or monitoring
CC progression or treatment of cancer, preferably lung or prostate cancer.
CC The present sequence represents the amino acid sequence of the human

transmembrane protein differentially expressed in prostate and lung
tumour TMPL

Sequence 490 AA:

Query Match 97.6%; Score 2294; DB 7; Length 490;
Best Local Similarity 100.0%; Pred. No. 8,4e-239;
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MESISMWSPKSLSTCLPNGINGIKDARKVTGVIGSGDPKSLTIRLRICGHHVIGS	60
DB	1	MESISMWSPKSLSTCLPNGINGIKDARKVTGVIGSGDPKSLTIRLRICGHHVIGS	60
QY	61	RNPKFASFPFPHVVDVTHHEDALTKTNIIFVALHREHYTSLMDLRHLVGVKILIDVSNM	120
DB	61	RNPKFASFPFPHVVDVTHHEDALTKTNIIFVALHREHYTSLMDLRHLVGVKILIDVSNM	120
QY	121	RINQYPSNAEYLALFPDSLIVKGNVVSAMALQGPDKASROYICSNNTQARQOYIE	180
DB	121	RINQYPSNAEYLALFPDSLIVKGNVVSAMALQGPDKASROYICSNNTQARQOYIE	180
QY	121	RINQYPSNAEYLALFPDSLIVKGNVVSAMALQGPDKASROYICSNNTQARQOYIE	180
DB	121	RINQYPSNAEYLALFPDSLIVKGNVVSAMALQGPDKASROYICSNNTQARQOYIE	180
QY	181	LARQNFPIIDGLSSAREINLPLRFTLMRGFVVAISLATFFLYSFVRDVIHPYA	240
DB	181	LARQNFPIIDGLSSAREINLPLRFTLMRGFVVAISLATFFLYSFVRDVIHPYA	240
QY	241	RNQSDFYKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQLYGTGKRRPPLMETLQ	300
DB	241	RNQSDFYKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQLYGTGKRRPPLMETLQ	300
QY	301	CRKQGLISFFPFAVVAAYSLCLPMRSERYLFLNMAVQOVHANIEENMEBEVRIEM	360
DB	301	CRKQGLISFFPFAVVAAYSLCLPMRSERYLFLNMAVQOVHANIEENMEBEVRIEM	360
QY	361	ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE	420
DB	361	ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE	420
QY	421	EEYRFPYPPNFVALVLPISIVL	444
DB	421	EEYRFPYPPNFVALVLPISIVL	444

RESULT 10

AAU10187 standard; protein; 490 AA.

AAU10187;

16-JAN-2002 (first entry)

Human Six-Transmembrane Protein of Prostate 1, STMP1.

Human; Six-Transmembrane Protein of Prostate 1; STM1; prostate cancer;
benign prostatic hyperplasia; acute prostatic disorder; lymphoma;
cryptorchidism; testicular disorder; proliferative disorder; breast cancer;
leukemia; melanoma; ovarian cancer; pancreatic cancer;
liver cancer; lung cancer; cytostatic.

Homo sapiens.

OS	Homo sapiens.
XX	
XX	Location/Qualifiers
FT	209..230
FT	/label=Transmembrane_domain_1
FT	255..273
FT	/label=Transmembrane_domain_2
FT	304..325
FT	/label=Transmembrane_domain_3
FT	360..380
FT	/label=Transmembrane_domain_4
FT	393..415
FT	/label=Transmembrane_domain_5
FT	432..452
FT	/label=Transmembrane_domain_6

XX MO200172962-A2.

XX 04-OCT-2001.

XX 23-MAR-2001; 2001MO-US0009410.

XX 24-MAR-2000; 2000US-0191929P.

XX (SAAT/) SAATCIGLU F.

XX Satcloglu F;

XX WPI, 2001-662926/76.

XX N-PSDB; AAS15793, AAS15801, AAS15802.

PT New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises
PT prostate-specific or testis-specific nucleic acids.

PS Claim 1; Fig 4B; 114pp; English.

XX The invention relates to substantially pure prostate-specific or testis-
XX specific polypeptides and the nucleic acids encoding them. Also included
XX are vectors and host cells expressing the proteins, a transgenic animal
XX expressing the protein, antibodies against the proteins, probes for
XX detecting the nucleic acids, antisense molecules for the nucleic acids
XX and methods of isolating modulators of the proteins. Compounds that
XX modulate the prostate specific or testis specific polypeptide are useful
XX to diagnose, prevent or treat disorders of the testis or prostate
XX particularly prostate cancer, benign prostatic hyperplasia, acute
XX prostaticitis, testicular cancer, cryptorchidism, undescended, retractile,
XX ascending or vanished testis. Other proliferative disorders for which the
XX modulators may be used include lymphoma, leukaemia, melanoma, ovarian
XX cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
XX The present sequence represents a prostate specific protein, Six-
XX Transmembrane Protein of Prostate 1, STM1.

XX Sequence 490 AA:

Query Match 97.4%; Score 2290; DB 4; Length 490;
Best Local Similarity 99.8%; Pred. No. 2.3e-238;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MESISMWSPKSLSTCLPNGINGIKDARKVTGVIGSGDPKSLTIRLRICGHHVIGS	60
DB	1	MESISMWSPKSLSTCLPNGINGIKDARKVTGVIGSGDPKSLTIRLRICGHHVIGS	60
QY	61	RNPKFASFPFPHVVDVTHHEDALTKTNIIFVALHREHYTSLMDLRHLVGVKILIDVSNM	120
DB	61	RNPKFASFPFPHVVDVTHHEDALTKTNIIFVALHREHYTSLMDLRHLVGVKILIDVSNM	120
QY	121	RINQYPSNAEYLALFPDSLIVKGNVVSAMALQGPDKASROYICSNNTQARQOYIE	180
DB	121	RINQYPSNAEYLALFPDSLIVKGNVVSAMALQGPDKASROYICSNNTQARQOYIE	180
QY	181	LARQNFPIIDGLSSAREINLPLRFTLMRGFVVAISLATFFLYSFVRDVIHPYA	240
DB	181	LARQNFPIIDGLSSAREINLPLRFTLMRGFVVAISLATFFLYSFVRDVIHPYA	240
QY	241	RNQSDFYKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQLYGTGKRRPPLMETLQ	300
DB	241	RNQSDFYKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQLYGTGKRRPPLMETLQ	300
QY	301	CRKQGLISFFPFAVVAAYSLCLPMRSERYLFLNMAVQOVHANIEENMEBEVRIEM	360
DB	301	CRKQGLISFFPFAVVAAYSLCLPMRSERYLFLNMAVQOVHANIEENMEBEVRIEM	360
QY	361	ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE	420
DB	361	ISFGIMSLGLSLAVTSIPSVSNALNWRFSFIQSTLGYVALLISTFHVLIYGMKRAFE	420
QY	421	EEYRFPYPPNFVALVLPISIVL	444

Db 421 EETRYFTPNFVLALVPSIVL 444

RESULT 11

AAU10189 standard; protein; 419 AA.

AAU10189;

16-JAN-2002 (first entry)

Human ORF3 of Six-Transmembrane Protein of Prostate 1.

Human; Six-Transmembrane Protein of Prostate 1; STM; 1; prostate cancer; benign prostatic hyperplasia; acute prostatitis; testicular cancer; cryptorchidism; testicular disorder; proliferative disorder; lymphoma; leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer; liver cancer; lung cancer; cytostatic; ORF3.

Homo sapiens.

MO200172962-A2.

04-OCT-2001.

23-MAR-2001; 2001MO-US009410.

24-MAR-2000; 2000US-0191929P.

(SAAT/) SAATCLOGU F.

Saatclogu F;

WPI; 2001-662926/76.

N-PSDB; AAS15813.

New polynucleotide for the diagnosis, prevention and treatment for prostate and testis disorders, particularly prostate cancer, comprises prostate-specific or testis-specific nucleic acids.

Claim 1; Fig 4K; 114pp; English.

The invention relates to substantially pure prostate-specific or testis-specific polypeptides and the nucleic acids encoding them. Also included are vectors and host cells expressing the proteins, a transgenic animal expressing the protein, antibodies against the proteins, probes for detecting the nucleic acids, antisense molecules for the nucleic acids and methods of isolating modulators of the proteins. Compounds that modulate the prostate specific or testis specific polypeptide are useful to diagnose, prevent or treat disorders of the testis or prostate particularly prostate cancer, benign prostatic hyperplasia, acute prostatitis, testicular cancer, cryptorchidism, undescended, retractile, ascending or vanished testis. Other proliferative disorders for which the modulators may be used include lymphoma, leukaemia, melanoma, ovarian cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer. The present sequence is prostate specific protein, Six-Transmembrane CC Protein of Prostate 1, STM1, ORF3

Sequence 419 AA;

Query Match 86.6%; Score 2036; DB 4; Length 419;

Best Local Similarity 99.7%; Pred. No. 5,5e-211; Indels 0; Gaps 0;

Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MESISMWSPKSLSECTLPNGINGIKDAKVTGVIGSGDPKSLTIRLRGYNVIGS 60

1 MESISMWSPKSLSECTLPNGINGIKDAKVTGVIGSGDPKSLTIRLRGYNVIGS 60

61 RNPKPSEFFPHVVDVTHHEDALTKNIIFAVIRHRYSLMDLRLHLLVGLKLLDVSNM 120

61 RNPKPSEFFPHVVDVTHHEDALTKNIIFAVIRHRYSLMDLRLHLLVGLKLLDVSNM 120

QY 121 RINQYPSNMEYLAISLPDLSLYKGFNVVSAMALQLGPKDASQVYICSNNTQARQVTE 180

Db 121 RINQYPSNMEYLAISLPDLSLYKGFNVVSAMALQLGPKDASQVYICSNNTQARQVTE 180

QY 181 LARQNFPIPDIGLSARREIENPLRLFTLMRGPVVAISLATFFLYGFVADVIHPYA 240

Db 181 LARQNFPIPDIGLSARREIENPLRLFTLMRGPVVAISLATFFLYGFVADVIHPYA 240

QY 241 RNOQSDFFKPIEIVNKTLPVIAITLLSLVYLAQLAAVQLYGGTKRRFPPLWLETWLQ 300

Db 241 RNOQSDFFKPIEIVNKTLPVIAITLLSLVYLAQLAAVQLYGGTKRRFPPLWLETWLQ 300

QY 301 CRKQGLISFFPFAVAVAYSLCLPMRSEERYLFNMAVQOVHANEENSMNEEYVRLEMY 360

Db 301 CRKQGLISFFPFAVAVAYSLCLPMRSEERYLFNMAVQOVHANEENSMNEEYVRLEMY 360

QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRBFSFIQ 395

Db 361 ISFGIMSLGLSLAVTSIPSVSNALNWRBFSFIQ 395

RESULT 12

ABG12306 standard; protein; 576 AA.

ABG12306;

18-FEB-2002 (first entry)

Novel human diagnostic protein #12297.

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001MO-US008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSB-) HYSBQ INC.

Dzmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

N-PSDB; AAS76493.

New isolated polynucleotide and encoded polypeptides, useful in PT diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 20; SEQ ID NO 42665; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 576 AA;

Query Match 83.9%; Score 1972.5; DB 4; Length 576;
 Best Local Similarity 87.4%; Pred. No. 6.7e-204;
 Matches 390; Conservative 0; Mismatches 1; Indels 55; Gaps 1;

QY 1 MESISMWGSPPKSLSETCLPENGINGIKDARKTVGVIGSGDFPKSLTIRLRGCHVIVIGS 60
 Db 1 MESISMWGSPPKSLSETCLPENGINGIKDARKTVGVIGSGDFPKSLTIRLRGCHVIVIGS 60
 QY 61 RNPKEASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 Db 61 RNPKEASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQYPSNAEYIASLPDLSLVKGFNVSAVALQGPDKASROYITCSNNIOAQOQVIE 180
 Db 121 RINQYPSNAEYIASLPDLSLVKGFNVSAVALQGPDKASROYITCSNNIOAQOQVIE 180
 QY 181 LAROLNFIPIIDGSLSSAREIENLPRLFTLRGPPVVAISLATFFFLYSFVADVHPYA 240
 Db 181 LAROLNFIPIIDGSLSSAREIENLPRLFTLRGPPVVAISLATFFFLYSFVADVHPYA 240
 QY 241 RNOQSDFYKIPRIEIVNKTLPPIVAITLLSVYAGLIAAAYOLYGTGKRPPFWLETWLQ 300
 Db 241 RNOQSDFYKIPRIEIVNKTLPPIVAITLLSVYAGLIAAAYOLYGTGKRPPFWLETWLQ 300
 QY 301 CRKQGLISFPFAMVVAISLCIPMRSEERYIFLMAVQOCHANIENSWNEEVRRIEMY 360
 Db 301 CRKQGLISFPFAMVVAISLCIPMRSEERYIFLMAVQOCHANIENSWNEEVRRIEMY 360
 QY 361 ISFGIMSLGLSLAIVTSIPVSNALNWRSEFISOSTIGYVALLISTFHVLIYGMKAAFE 420
 Db 361 ISFGIMSLGLSLAIVTSIPVSNALNWRSEFISOSTIGYVALLISTFHVLIYGMKAAFE 420
 QY 421 EYYRFPYTPPNFVLAIVLPSIVLIDL 446
 Db 421 EYYRFPYTPPNFVLAIVLPSIVLIDL 391

RESULT 13

ABG00113
 ID ABG00113 standard, protein; 1273 AA.

XX ABG00113;
 AC 13-FEB-2002 (first entry)
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #104.
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM Food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX

PI Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI; 2001-639362/73.
 DR N-PDB; AAS64300.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 30472; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridization probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping;
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 1273 AA;

Query Match 83.9%; Score 1972.5; DB 4; Length 1273;
 Best Local Similarity 87.4%; Pred. No. 2.3e-203;
 Matches 390; Conservative 0; Mismatches 1; Indels 55; Gaps 1;

QY 1 MESISMWGSPPKSLSETCLPENGINGIKDARKTVGVIGSGDFPKSLTIRLRGCHVIVIGS 60
 Db 73 MESISMWGSPPKSLSETCLPENGINGIKDARKTVGVIGSGDFPKSLTIRLRGCHVIVIGS 132
 QY 61 RNPKEASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 Db 133 RNPKEASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 192
 QY 121 RINQYPSNAEYIASLPDLSLVKGFNVSAVALQGPDKASROYITCSNNIOAQOQVIE 180
 Db 121 RINQYPSNAEYIASLPDLSLVKGFNVSAVALQGPDKASROYITCSNNIOAQOQVIE 252
 QY 181 LAROLNFIPIIDGSLSSAREIENLPRLFTLRGPPVVAISLATFFFLYSFVADVHPYA 240
 Db 253 LAROLNFIPIIDGSLSSAREIENLPRLFTLRGPPVVAISLATFFFLYSFVADVHPYA 312
 QY 241 RNOQSDFYKIPRIEIVNKTLPPIVAITLLSVYAGLIAAAYOLYGTGKRPPFWLETWLQ 300
 Db 313 RNOQSDFYKIPRIEIVNKTLPPIVAITLLSVYAGLIAAAYOLYGTGKRPPFWLETWLQ 372
 QY 301 CRKQGLISFPFAMVVAISLCIPMRSEERYIFLMAVQOCHANIENSWNEEVRRIEMY 360
 Db 373 CRKQGLISFPFAMVVAISLCIPMRSEERYIFLMAVQOCHANIENSWNEEVRRIEMY 411
 QY 361 ISFGIMSLGLSLAIVTSIPVSNALNWRSEFISOSTIGYVALLISTFHVLIYGMKAAFE 420
 Db 412 ISFGIMSLGLSLAIVTSIPVSNALNWRSEFISOSTIGYVALLISTFHVLIYGMKAAFE 437
 QY 421 EYYRFPYTPPNFVLAIVLPSIVLIDL 446
 Db 438 EYYRFPYTPPNFVLAIVLPSIVLIDL 463

RESULT 14
AA02841
ID AAE02841 standard; protein: 450 AA.
XX AAE02841;
AC
XX 06-AUG-2001 (first entry)
XX
XX Human STEAP-2 protein, alternative version.
DE
XX Human, cytototoxic; antiproliferative; vaccine; gene therapy;
KM six transmembrane epithelial antigen of the prostate-2; STEAP-2;
KM chromosome 7q21; cancer; prostate; colon; bladder; lung; ovarian;
KM pancreatic.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Misc-difference 335..336
FT /note= "Encoded by GACTGAGCT"
FT Misc-difference 375..376
FT /note= "Encoded by AGATGAGCT"
FT Misc-difference 415..416
FT /note= "Encoded by GAGTAAAGC"
FT Misc-difference 445..446
FT /note= "Encoded by ACATGAGCT"
FT Misc-difference 448..449
FT /note= "Encoded by AATTAAATTC"
XX
XX WO200140276-A2.
XX
XX 07-JUN-2001.
XX
XX 06-DEC-2000; 2000WO-US033040.
XX
XX 06-DEC-1999; 99US-00455486.
XX
XX (UROG-) UROGENESYS INC.
XX
XX Afar DH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
PI Paris W, Jakobovits A;
XX
XX WPI: 2001-367804/38.
XX
XX N-PSDB; AAD07072.
XX
XX New STEAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX
XX Claim 1; Page 175-176; 187pp; English.
XX
XX The present sequence is an alternative version of human six transmembrane
XX epithelial antigen of the prostate (STEAP)-2 protein. STEAP-2 gene is a member
XX of cell surface serpentine transmembrane antigens. STEAP-2 gene is
XX located on chromosome 7q21 and is used in gene therapy. Inhibiting the
XX development or progression of a cancer (eg. prostate, colon, bladder,
XX lung, ovarian and pancreatic) expressing STEAP or inhibiting growth or
XX killing cells expressing STEAP in a patient, comprises administering a
XX vaccine composition to the patient. Treating a patient with a cancer that
XX expresses STEAP, or inhibiting growth or killing cells expressing STEAP,
XX comprises administering to the patient a vector encoding single chain
XX monoclonal antibody that comprises the variable domains of the heavy and
XX light chains of the monoclonal antibody that specifically binds to STEAP,
XX such that the vector delivers the single chain monoclonal antibody coding
XX sequence to the cancer cells and the encoded single chain monoclonal
XX antibody is expressed intracellularly. Note: This sequence is stated to
XX be the same as that being shown as SEQ ID NO:8 (AA02781) in figure 9A-9C
XX of the specification. However the present sequence lacks several amino
XX acids at its N-terminal end and has additional amino acids at its C-
XX terminal end when compared with the sequence shown in figure 9A-9C
XX
XX Sequence 450 AA;
XX
XX

Query Match 73.9%; Score 1738; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 1,1e-178;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 119 NMRINOYVESNAEYLAFLPDSLIVKGFNVSAWALQGPASROYICSNNOIARQOV 178
DB 1 NMRINOYVESNAEYLAFLPDSLIVKGFNVSAWALQGPASROYICSNNOIARQOV 60
QY 179 IELARQNLFPIDIGLSASAREIENLPLRFTLRGQVVAISLATFPFLYSPFVDYIHP 238
DB 61 IELARQNLFPIDIGLSASAREIENLPLRFTLRGQVVAISLATFPFLYSPFVDYIHP 120
QY 239 YARNQSPFYKIPLEIVNKTLPVATLLSLVYAGLAAAYOLYGGKRRFPMTETW 298
DB 121 YARNQSPFYKIPLEIVNKTLPVATLLSLVYAGLAAAYOLYGGKRRFPMTETW 180
QY 239 LQCRKQLGLSFFPAMVAVYSLCLPMRSERYFLNNAVQOVHANINENWNEEVRLE 358
DB 181 LQCRKQLGLSFFPAMVAVYSLCLPMRSERYFLNNAVQOVHANINENWNEEVRLE 240
QY 359 MYISFGIMSLGLSLAVTSIPSVSNALNRRFSGIOTIGVALLISTFHVLYGMKRA 418
DB 241 MYISFGIMSLGLSLAVTSIPSVSNALNRRFSGIOTIGVALLISTFHVLYGMKRA 300
QY 419 FEEYFRFTPPNFVLTALVPSIVLIDLLQCRYPD 454
DB 301 FEEYFRFTPPNFVLTALVPSIVLIDLLQCRYPD 336
RESULT 15
ABP62883
ID ABP62883 standard; protein: 1082 AA.
XX
XX ABP62883;
XX
XX 14-OCT-2002 (first entry)
XX
XX Human polypeptide SEQ ID NO 320.
XX
XX Human; vulnereary; dermatological; neuroprotective; nootropic; cancer;
XX antiparkinsonian; immunostimulant; cytostatic; immunosuppressive;
XX antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
XX burn; central nervous system disorder; Alzheimer's disease;
XX Parkinson's disease; Huntington's disease; immune disorder;
XX autoimmune disorder; multiple sclerosis; diabetes; allergy.
XX
XX Homo sapiens.
XX
XX WO200218424-A2.
XX
XX 07-MAR-2002.
XX
XX 31-AUG-2001; 2001WO-US027093.
XX
XX 01-SEP-2000; 2000US-00654935.
XX
XX (HYSR-) HYSR INC.
XX
XX Tang YT, Auandi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
PI Zhao QA, Wang D, Liu C, Dymnac RT, Wehrman T;
XX
XX WPI: 2002-583321/62.
XX
XX N-PSDB; ABQ93362.
XX
XX New polynucleotide and polypeptides, useful for treatment and diagnosis
XX of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
XX sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
XX sclerosis, diabetes and allergies.
XX
XX Claim 20; SEQ ID NO 320; 284pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (i) comprising one of
XX 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising

CC administering to a mammalian subject a composition comprising the proteol-
 CC (II) encoded by (I) (ABP62809-ABP6305) or an antibody (III) to (II).
 CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.
 CC (I) is useful for gene therapy of diseases and (II) can be used for
 CC therapeutic treatment. Diseases that may be treated include wound healing
 CC and tissue repair, burns, central nervous system disorders (e.g.
 CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
 CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
 CC sclerosis, diabetes and allergies. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WFO at
 CC ftp.wipo.int/pub/published_pat_sequences

SQ Sequence 1082 AA;

Query Match	58.0%	Score	1364.5	-DB 5	Length	1082
Best Local Similarity	83.2%	Pred. No.	1.1e-137			
Matches 272	Conservative	0	Mismatches	0	Indels	55
					Gaps	1

QY	120	MRINOVESNAEYLAISFPBSLIVKGFNVWSMAUOLPXPASQOYICSNIOARQOVI	179
Db	1	MRINOVESNAEYLAISFPBSLIVKGFNVWSMAUOLPXPASQOYICSNIOARQOVI	60
QY	180	ELAROLNFIPIIDVGSJSSABEINBLRFTLMRGPVVAISLATFFPLSPFADVIHPI	239
Db	61	ELAROLNFIPIIDVGSJSSABEINBLRFTLMRGPVVAISLATFFPLSPFADVIHPI	120
QY	240	ARNOQSDFYKPIPIEIVNKTLPVAILTSLVYLAGLAAAYOLYGGTKYRRFPFMEIETWL	299
Db	121	ARNOQSDFYKPIPIEIVNKTLPVAILTSLVYLAGLAAAYOLYGGTKYRRFPFMEIETWL	180
QY	300	QCRKOLGLJSPFFPAMNHVAVSLCLPMRBSERLIFLMAAYQOVHANIENSMNEEYVMIEM	359
Db	181	QCRKOLGLJSPFFPAMNHVAVSLCLPMRBSERLIFLMAAYQOVHANIENSMNEEYVMIEM	220
QY	360	YISFGIMSLGLSHLAVTISIPVSANLMBREESFIOSTLGYVALHISTFHVLIYGMRAF	419
Db	221	YISFGIMSLGLSHLAVTISIPVSANLMBREESFIOSTLGYVALHISTFHVLIYGMRAF	245
QY	420	EEERYREYTPBNFVLAIVLPSIVITDL	446
Db	246	EEERYREYTPBNFVLAIVLPSIVITDL	272

Search completed: April 16, 2004, 12:39:11
Job time : 63 secs

QY 121 RINQPESSNAEYLAIFPDSLIVKGFNVSAWALQGPXDSROVYICSNIIQARQVIE 180
 DB 121 RINQPESSNAEYLAIFPDSLIVKGFNVSAWALQGPXDSROVYICSNIIQARQVIE 180
 QY 181 LARQNLFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240
 DB 181 LARQNLFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240
 QY 241 RNOQSDPYKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQYGTXYRRPFPMLETWQ 300
 DB 241 RNOQSDPYKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQYGTXYRRPFPMLETWQ 300
 QY 301 CRKQGLISFFFAWVAVASLCLEPMSRSELYFLNMAVYQVHANINENSMBEEMRIEMX 360
 DB 301 CRKQGLISFFFAWVAVASLCLEPMSRSELYFLNMAVYQVHANINENSMBEEMRIEMX 360
 QY 361 ISFGIMSLGSLSLAVTSIPSVNALNMRBFSFIQSTLGYVALLISTFPHLYIGMKRAFE 420
 DB 361 ISFGIMSLGSLSLAVTSIPSVNALNMRBFSFIQSTLGYVALLISTFPHLYIGMKRAFE 420
 QY 421 EBYRYFTPPNPFVATLVPISVIL 444
 DB 421 EBYRYFTPPNPFVATLVPISVIL 444

RESULT 2

ID Q8BFT2 PRELIMINARY; PRT; 490 AA.
 AC Q8BFT2;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Six-transmembrane epithelial antigen of prostate 2.
 GN STAP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RA Poikka K.P., Helenius M.A., Visakorpi T.;
 RT "Cloning and characterization of a novel six-transmembrane protein
 STAP2, down-regulated in androgen-independent prostate cancer.";
 RL Submitted (DSC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF45138; AA04080.1;
 DR Genew; HGNC:17885; STAP2.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR004455; NADPoxred.F420.
 DR Pfam; PF03807; F420_oxidored; I.
 DR KX
 KW Transmembrane.
 SQ SEQUENCE 490 AA; 55961 MW; F2B9C30CDACEF81 CRC64;

Query Match 97.2%; Score 2284; DB 4; Length 490;
 Best Local Similarity 99.8%; Pred. No. 1.2e-175;
 Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISMGSPPKSLSTCTCPNGINGIKDKARKYTVGVIGSGDFAKSLTIRLIRCGYHVIGS 60
 DB 1 MESISMGSPPKSLSTCTCPNGINGIKDKARKYTVGVIGSGDFAKSLTIRLIRCGYHVIGS 60
 QY 61 RNPKEASEFPFHVVDVTHEDALTNTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHEDALTNTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQPESSNAEYLAIFPDSLIVKGFNVSAWALQGPXDSROVYICSNIIQARQVIE 180
 DB 121 RINQPESSNAEYLAIFPDSLIVKGFNVSAWALQGPXDSROVYICSNIIQARQVIE 180
 QY 181 LARQNLFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240
 DB 181 LARQNLFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240

DB 181 LARQNLFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240
 QY 241 RNOQSDPYKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQYGTXYRRPFPMLETWQ 300
 DB 241 RNOQSDPYKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQYGTXYRRPFPMLETWQ 300
 QY 301 CRKQGLISFFFAWVAVASLCLEPMSRSELYFLNMAVYQVHANINENSMBEEMRIEMX 360
 DB 301 CRKQGLISFFFAWVAVASLCLEPMSRSELYFLNMAVYQVHANINENSMBEEMRIEMX 360
 QY 361 ISFGIMSLGSLSLAVTSIPSVNALNMRBFSFIQSTLGYVALLISTFPHLYIGMKRAFE 420
 DB 361 ISFGIMSLGSLSLAVTSIPSVNALNMRBFSFIQSTLGYVALLISTFPHLYIGMKRAFE 420
 QY 421 EBYRYFTPPNPFVATLVPISVIL 444
 DB 421 EBYRYFTPPNPFVATLVPISVIL 444

RESULT 3

ID Q8BWB6 PRELIMINARY; PRT; 489 AA.
 AC Q8BWB6;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Weakly similar to tumor suppressor p19DE.
 GN 4921538B1/RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RC MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium;
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK052981; BAC35230.1;
 DR MGI; MGI:1521301; 4921538B1/RIK.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR004455; NADPoxred.F420.
 DR Pfam; PF03807; F420_oxidored; I.
 SQ SEQUENCE 489 AA; 55760 MW; 98CD63D59DDDF24C CRC64;

Query Match 95.3%; Score 2239.5; DB 11; Length 489;
 Best Local Similarity 97.3%; Pred. No. 4.8e-172;
 Matches 432; Conservative 9; Mismatches 2; Indels 1; Gaps 1;

QY 1 MESISMGSPPKSLSTCTCPNGINGIKDKARKYTVGVIGSGDFAKSLTIRLIRCGYHVIGS 60
 DB 1 MESISMGSPPKSLSTCTCPNGINGIKDKARKYTVGVIGSGDFAKSLTIRLIRCGYHVIGS 60
 QY 61 RNPKEASEFPFHVVDVTHEDALTNTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHEDALTNTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQPESSNAEYLAIFPDSLIVKGFNVSAWALQGPXDSROVYICSNIIQARQVIE 180
 DB 121 RINQPESSNAEYLAIFPDSLIVKGFNVSAWALQGPXDSROVYICSNIIQARQVIE 180
 QY 181 LARQNLFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240
 DB 181 LARQNLFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240
 QY 241 RNOQSDPYKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQYGTXYRRPFPMLETWQ 300
 DB 241 RNOQSDPYKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQYGTXYRRPFPMLETWQ 300
 QY 301 CRKQGLISFFFAWVAVASLCLEPMSRSELYFLNMAVYQVHANINENSMBEEMRIEMX 360
 DB 301 CRKQGLISFFFAWVAVASLCLEPMSRSELYFLNMAVYQVHANINENSMBEEMRIEMX 360

```

Db      300 CRKQGLSLFFFAVVAHSLCLPMRSEHYFLNVAIQCHNINANNEEVRINRY 359
Qy      361 ISFGIMSLGLSLAVTSIPSVNALNMRBFSIQSTGLVALLSTFHYLYGWRKPE 420
Db      360 ISFGIMSLGLSLAVTSIPSVNALNMRBFSIQSTGLVALLSTFHYLYGWRKPE 419
Qy      421 EERYRPTPNFVLAIVLPSIVIL 444
Db      420 EERYRPTPNFVLAIVLPSIVIL 443

RESULT 4
Q8C5F0 PRELIMINARY; PRT; 488 AA.
AC Q8C5F0;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE DUDUIN 2 homolog.
GN TSAP6 OR 1010001D0IRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RL EMBL: AK078769; BAC73783.1;
DR MGD; MGI:1915678; Tsap6.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
SQ SEQUENCE 488 AA; 54780 MW; FDfCAf42AE503D11 CRC64;

Query Match 56.1%; Score 1318; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 8.4e-98;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

Qy      6 MMGSPKSLSETCLPENGINGIKDKARKVTGVYIGSGDFPAKSLTIRLRGCHVIVGSRNPKF 65
Db      14 LVDSGSLAE--VP-----KEAPK--VGILSGDFPARSLATRLVSGSFVVGSRNPKR 63
Qy      66 ASEFPFHVVDVTHEDALTKNIIIFVAIHREHYTSLMDLRHLVSKLIDVSNMRIN-- 123
Db      64 TAGLFPFLAQVTFQEEANMSPPEVITFAVFRHYSSLCGLADQLAGKILVDSNPTEKRL 123
Qy      124 QYPSNAEYLAFLPDSLIIVKGFNVSAWALQGPDKASROYICSNNIQARQYIELAR 183
Db      124 QHRQSNAEYLAFLPACTVVAFAFVISAALQAGPRDGRVLLICDQPAKRTISMAR 183
Qy      184 QLNFTPIDGLSSARETENIPLRLFTLMRGVVAALSLATFFFLYSFVVDVHPYARNO 243
Db      184 AMGFPTPLDMGSLASAREVEALPLRLPSWKVFTLLALGFLVCFTYVFIIDVLOPYIRKD 243
Qy      244 QSDFYKIPLEIVNKTLPVIAITLLSLVYLAGLAAAYQLYGTXYRRFPPLMETWLOCRK 303
Db      244 ENKRYKPLSVNTTLPVAVYLLSLVYLPGLVLAALQLRGRTKXQRFPMIDHMLQHRK 303
Qy      304 QGLISFPFAMHVAYSICLPMRSEHYFLNMAIQCHNINANNEEVRINRYISF 363
Db      304 QIGLISFPFAMHVAYSICLPMRSEHYFLNMAIQCHNINANNEEVRINRYISL 363
Qy      364 GIMSGLSLAVTSIPSVNALNMRBFSIQSTGLVALLSTFHYLYGWRKAFEEY 423
Db      364 GYLAIGMSLAVTSIPSVNALNMRBFSIQSTGLVALLSTFHYLYGWRKAFEEY 423

```

```

Qy      424 YRFYTPNFVLAIVLPSIVIL 444
Db      424 YRFYTPNFVLAIVLPSIVIL 444

RESULT 5
Q99P41 PRELIMINARY; PRT; 488 AA.
AC Q99P41;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Tumor suppressor pHyde.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunning; TISSUE=Prostatic carcinoma;
RX MEDLINE=20424188; PubMed=10969787;
RA Steiner M.S., Zhang X., Wang Y., Lu Y.;
RT "Growth inhibition of prostate cancer by an adenovirus expressing a
RT novel tumor suppressor gene, pHyde."
RL Cancer Res. 60:4419-4425(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunning; TISSUE=Prostatic carcinoma;
RA Lu Y., Rinaldy A.R., Steiner M.S.;
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Steiner M.S., Rinaldy A.;
RT "A novel prostate cancer associated gene."
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
RL EMBL: AF335281; AAK00361.1;
DR EMBL; AF338865; AAL78207.1;
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; 1.
SQ SEQUENCE 488 AA; 54640 MW; 59FF07121919FDC8 CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 1.2e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

Qy      6 MMGSPKSLSETCLPENGINGIKDKARKVTGVYIGSGDFPAKSLTIRLRGCHVIVGSRNPKF 65
Db      14 LVDSGSLAE--VP-----KEAPK--VGILSGDFPARSLATRLVSGSFVVGSRNPKR 63
Qy      66 ASEFPFHVVDVTHEDALTKNIIIFVAIHREHYTSLMDLRHLVSKLIDVSNMRIN-- 123
Db      64 TAGLFPFLAQVTFQEEANMSPPEVITFAVFRHYSSLCGLADQLAGKILVDSNPTEKRL 123
Qy      124 QYPSNAEYLAFLPDSLIIVKGFNVSAWALQGPDKASROYICSNNIQARQYIELAR 183
Db      124 QHRQSNAEYLAFLPACTVVAFAFVISAALQAGPRDGRVLLICDQPAKRTISMAR 183
Qy      184 QLNFTPIDGLSSARETENIPLRLFTLMRGVVAALSLATFFFLYSFVVDVHPYARNO 243
Db      184 AMGFPTPLDMGSLASAREVEALPLRLPSWKVFTLLALGFLVCFTYVFIIDVLOPYIRKD 243
Qy      244 QSDFYKIPLEIVNKTLPVIAITLLSLVYLAGLAAAYQLYGTXYRRFPPLMETWLOCRK 303
Db      244 ENKRYKPLSVNTTLPVAVYLLSLVYLPGLVLAALQLRGRTKXQRFPMIDHMLQHRK 303
Qy      304 QGLISFPFAMHVAYSICLPMRSEHYFLNMAIQCHNINANNEEVRINRYISF 363
Db      304 QIGLISFPFAMHVAYSICLPMRSEHYFLNMAIQCHNINANNEEVRINRYISL 363
Qy      364 GIMSGLSLAVTSIPSVNALNMRBFSIQSTGLVALLSTFHYLYGWRKAFEEY 423
Db      364 GYLAIGMSLAVTSIPSVNALNMRBFSIQSTGLVALLSTFHYLYGWRKAFEEY 423

```

QY 424 YRFYTPNFVTLALVPSIVIL 444
 DB 424 YKFYLPFTFTLLPVCVIL 444

RESULT 6

Q8C159 PRELIMINARY; PRT; 488 AA.
 AC Q8C159;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Similar to RIKEN cDNA 1010001D01 gene.
 GN TSAP6 OR 1010001D01RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J
 RA Strauberg R.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC037435; AAH37435.1; -.
 DR MGD; MGI:1915678; Tmap6.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420_oxidored; 1.
 DR SEQUENCE 488 AA; 54749 MW; 9A08D99C9C83F4 CRC64;
 SQ

Query Match 56.0%; Score 1316; DB 11; Length 488;
 Best Local Similarity 56.2%; Pred. No. 1.2e-97;
 Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPENGINGIDARKVTYGVIGSGDFAKSLTIRLCGYHVIIGSRNPKF 65
 DB 14 LVSDSGSLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSFVVGSRNPKR 63
 QY 66 ASEFPFHYVDVTHEDALTKNIIIFVAIHREHYTSLMDLHILVSKILLIDVSNMNRIN-- 123
 DB 64 TAGLEPSLAQVTPQSEAVSPEVIFVAFVREHYSSLSLADQLAGKILLVDVSNTEKEHL 123
 QY 124 QPESNAEYLASLPDLSILVGFNVVSAWALQGPCKASROYVICSNNIQARQVIELAR 183
 DB 124 QHRQSNAEYLASLPDLSILVGFNVVSAWALQGPCKASROYVICSNNIQARQVIELAR 183
 QY 184 QLNFPIDIGSLSSARETENPLRLFTLMRGVVAISLATFFFLYSFVRDVIHPYARNQ 243
 DB 184 AMGFPLDMGSLASREVEALPLRLPSWKVPTLLALGLFVCFYTYNPIRDVLQPIYRKD 243
 QY 244 QSDFKYPIEIVNKTLPVATITLLSLVYLAGLLAAAYQVYGYKRRFPPLWTMLQCRK 303
 DB 244 ENKFFYKMPVNTLLPCVAVYLLSLVYLPGLVLAALQLRGTXYQRPDMLDHWLQHRK 303
 QY 304 QGLISLSPFAMVAVYSLCPMRSEERYLFNMAVQOVHANIENSNMNEEYVRIEMYSF 363
 DB 304 QIGLSLSPFAMVAVYSLCPMRSEERYLFNMAVQOVHANIENSNMNEEYVRIEMYSF 363
 QY 364 GIMSGLSLAVTSIPSVSNALNWRBFSFIQSTLGYVALLISTFHVLIYGMKRAFEERY 423
 DB 364 GYVALGMSLAVTSIPSVSNALNWRBFSFIQSTLGYVALLISTFHVLIYGMKRAFEERY 423
 QY 424 YRFYTPNFVTLALVPSIVIL 444
 DB 424 YKFYLPFTFTLLPVCVIL 444

RESULT 7

Q8OZF3 PRELIMINARY; PRT; 526 AA.
 AC Q8OZF3;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Tsap6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSMZ ACC331; TISSUE=Bone marrow;
 RX MEDLINE=22506415; PubMed=12606722;
 RA Passer B.U., Nancy-Portebois V., Amzallag N., Prieur S., Cane C.,
 RA Roborel de Climens A., Finucci G., Bouvard V., Tunder M., Sushni L.,
 RA Morchoisne S.P., Cribble V., Lespagnol A., Dausset J., Oren M.,
 RA Amson R., Teichmann A.;
 RT "The p53-inducible Tsap6 gene product regulates apoptosis and the cell
 cycle and interacts with Nix and the Myl1 kinase."
 RT Proc. Natl. Acad. Sci. U.S.A. 100:2284-2289(2003).
 DR EMBL; AY214462; AA038239.1; -.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR004455; NADPoxred_F420.
 DR Pfam; PF03807; F420_oxidored; 1.
 DR SEQUENCE 526 AA; 58530 MW; 6306CD717E25200A CRC64;
 SQ

Query Match 56.0%; Score 1316; DB 11; Length 526;
 Best Local Similarity 56.2%; Pred. No. 1.3e-97;
 Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPENGINGIDARKVTYGVIGSGDFAKSLTIRLCGYHVIIGSRNPKF 65
 DB 52 LVSDSGSLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSFVVGSRNPKR 101
 QY 66 ASEFPFHYVDVTHEDALTKNIIIFVAIHREHYTSLMDLHILVSKILLIDVSNMNRIN-- 123
 DB 102 TAGLEPSLAQVTPQSEAVSPEVIFVAFVREHYSSLSLADQLAGKILLVDVSNTEKEHL 161
 QY 124 QPESNAEYLASLPDLSILVGFNVVSAWALQGPCKASROYVICSNNIQARQVIELAR 183
 DB 162 QHRQSNAEYLASLPDLSILVGFNVVSAWALQGPCKASROYVICSNNIQARQVIELAR 221
 QY 184 QLNFPIDIGSLSSARETENPLRLFTLMRGVVAISLATFFFLYSFVRDVIHPYARNQ 243
 DB 222 AMGFPLDMGSLASREVEALPLRLPSWKVPTLLALGLFVCFYTYNPIRDVLQPIYRKD 281
 QY 244 QSDFKYPIEIVNKTLPVATITLLSLVYLAGLLAAAYQVYGYKRRFPPLWTMLQCRK 303
 DB 282 ENKFFYKMPVNTLLPCVAVYLLSLVYLPGLVLAALQLRGTXYQRPDMLDHWLQHRK 341
 QY 304 QGLISLSPFAMVAVYSLCPMRSEERYLFNMAVQOVHANIENSNMNEEYVRIEMYSF 363
 DB 342 QIGLSLSPFAMVAVYSLCPMRSEERYLFNMAVQOVHANIENSNMNEEYVRIEMYSF 401
 QY 364 GIMSGLSLAVTSIPSVSNALNWRBFSFIQSTLGYVALLISTFHVLIYGMKRAFEERY 423
 DB 402 GYVALGMSLAVTSIPSVSNALNWRBFSFIQSTLGYVALLISTFHVLIYGMKRAFEERY 461
 QY 424 YRFYTPNFVTLALVPSIVIL 444
 DB 462 YKFYLPFTFTLLPVCVIL 482

RESULT 8

Q924Z1 PRELIMINARY; PRT; 514 AA.
 AC Q924Z1;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE DudaLin 2.
 GN TSAP6 OR 1010001D01RIK.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;
 [1]
 RA SEQUENCE FROM N.A.
 RA Serru V., Lambdin D., Lenoir C., Manivet P., Vaudourdolle M.,
 RA Kellermann O., Loric S.,
 RA "Molecular cloning and expression of mouse dudulin 2,"
 RA Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY029586; AAK50539.1;
 DR MGI: 1915678; Tsad6.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR InterPro: IPR004455; NADPoxred.F420.
 DR Pfam: PF03807; F420_oxidored; I.
 DR PROSITE: PS00290; IG_MHC; 1.
 SQ SEQUENCE 514 AA; 57268 MW; 339886C289AEC0B2 CRC64;
 Query Match 55.6%; Score 1306; DB 11; Length 514;
 Best Local Similarity 56.2%; Pred. No. 8.3e-97;
 Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;
 QY 6 MMSGPKSLSTCPLNGINGIKDARKVTGVIGSGDFAKSLTITRLRCGHVYIGSNPKF 65
 DB 14 LVSDSGSLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSGFVVGSHNPKR 63
 QY 66 ASEFPFVVDVTHEDALTKNTIIFVAIHREHYTSLMDLRLHVLGKILIDVSNMKN-- 123
 DB 64 TAGLFBELAQVTPOEAVSPEVIFVAVFREHYSLCSLADQLAGKILIDVSNPTEKHL 123
 QY 124 QYSESNAYIASLPDLSLVKGNVVSAMALQLGPKDASRGVYICSNNTIQAQOVTELAR 183
 DB 124 QHRSNAYIASLPACTVKAFFNVSAMALQAVPRDGNQVILICNDSSKAKORVEMAR 183
 QY 184 QLPFIPIDLSISAREIENLPLRLFTLMRGPVVAISLATFEFLYSPVDVHPFARQ 243
 DB 184 AMGTFPLDWSLSAREVEAIPRLDPSGVPTLLAGLFCVCTYTNFTIDVQPIRKD 243
 QY 244 QSDPYKPIPIEIVNKTLPVATITLSLVYLAGLLAAAYLYGKRRFPEMTLQCR 303
 DB 244 ENKRYKMPLEVVNTLPCVAYVLLSLVYLPGLVLAALQLRGTKYQRPDMLDHLQCR 303
 QY 304 QLGILSFFPAMVAVYSLCPMRSEERYLFLNMAQOVHANINSMNEEVRMIEMVIF 363
 DB 304 QIGLSFFPMLHALYFCLPRLRSRDYDVLNIAKVLANKRLMVEEVMMEIYLSL 363
 QY 364 GINSIGLSLAVTISPSVSNALNWRSEFISQSTLGVALISTFHVLIYMKRAPEEB 423
 DB 364 GVALGMLSLAVSLPSIANSINMKSEFVQSLGFVALISTFHVLTLYGWTAREEH 423
 QY 424 YRFTPPNPLVALVPSIVL 444
 DB 424 YKFLPPTFTLTLLPCVIL 444
 RESULT 9
 Q86SF6 PRELIMINARY; PRT; 488 AA.
 AC Q86SF6; (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Similar to dudulin 2 (TSAd6).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Kidney;
 RA Strausberg R.;
 RU Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=22506415; PubMed=12606722;
 RA Passer B.J., Nancy-Portebois V., Amzallig N., Priet S., Gans C.,
 RA Roborel de Climens A., Fituci G., Bouvard V., Tuyen M., Susini L.,

Morchoisne S.P., Cribble V., Iespagnol A., Dausset J., Oren M.,
 RA Amson R., Telemann A.;
 RA "The p53-inducible TSAd6 gene product regulates apoptosis and the cell
 RT cycle and interacts with Nix and the Myt1 kinase,"
 RT Proc. Natl. Acad. Sci. U.S.A. 100:2284-2289(2003).
 DR EMBL: BC042150; AAA42150.1;
 DR EMBL: AY214461; AAC38238.1;
 DR GO: GO:0006118; P:electron transport; IEA.
 DR InterPro: IPR003006; IG_MHC.
 DR InterPro: IPR004455; NADPoxred.F420.
 DR Pfam: PF03807; F420_oxidored; I.
 DR PROSITE: PS00290; IG_MHC; 1.
 SQ SEQUENCE 488 AA; 54600 MW; C09EB0D0430F9BFB CRC64;
 Query Match 54.1%; Score 1272; DB 4; Length 488;
 Best Local Similarity 54.5%; Pred. No. 4.3e-94;
 Matches 242; Conservative 82; Mismatches 108; Indels 12; Gaps 4;
 QY 3 SISMGSPKSLSTCPLNGINGIKDARKVTGVIGSGDFAKSLTITRLRCGHVYIGSN 62
 DB 11 SLHLDSSSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGSGFVVGSHN 60
 QY 63 PFXASEFPFVVDVTHEDALTKNTIIFVAIHREHYTSLMDLRLHVLGKILIDVSN--NM 120
 DB 61 PXTARLPFSAQVTPOEAVSPEVIFVAVFREHYSLCSLADQLAGKILIDVSNPTEQ 120
 QY 121 RINQVPSNAYIASLPDLSLVKGNVVSAMALQLGPKDASRGVYICSNNTIQAQOVIE 180
 DB 121 EHLQRESNAYIASLPFTCTVKAFFNVSAMTLOAGRDGNQVILICNDSSKAKORVEMAR 180
 QY 181 LARQNFIPIDLSISAREIENLPLRLFTLMRGPVVAISLATFEFLYSPVDVHPFARQ 240
 DB 181 MALAMGFVPMVMSGLASMEVEAMPRLPLPAKVPITLLAGLFCVCFAYANFVRDVLQPYV 240
 QY 241 RNQSDPYKPIPIEIVNKTLPVATITLSLVYLAGLLAAAYLYGKRRFPEMTLQCR 300
 DB 241 QESQKPFKLPVSVNTLPCVAYVLLSLVYLPGLVLAALQLRGTKYQRPDMLDHLQ 300
 QY 301 CRKQGLSFFPAMVAVYSLCPMRSEERYLFLNMAQOVHANINSMNEEVRMIEMVIF 360
 DB 301 HRKQGLSFFPAMVAVYSLCPMRSEERYLFLNMAQOVHANINSMNEEVRMIEMVIF 360
 QY 361 ISFGMSIGLSLAVTISPSVSNALNWRSEFISQSTLGVALISTFHVLIYMKRAPE 420
 DB 361 ISLGVIALGTLAVTISPSIANSINMKSEFVQSLGFVALISTFHVLTLYGWTARE 420
 QY 421 ERYRFPNPLVALVPSIVL 444
 DB 421 ERYRFPNPLVALVPSIVL 444
 RESULT 10
 Q72389 PRELIMINARY; PRT; 498 AA.
 AC Q72389;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein DKFZP686H07150.
 GN DKFZP686H07150.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Human colon endocel primary cell culture;
 RA Bloeker H., Boecher M., Mewes H.W., Weil B., Amid C., Oeanger A.,
 RA Robo G., Han M., Wiemann S.;
 RU Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BX538047; CND97986.1;
 KW Hypothetical protein.
 SQ SEQUENCE 498 AA; 55593 MW; 69DD23D5A00C2D6B CRC64;

Query Match	54.1%	Score 1272	DB 4	Length 498
Best Local Similarity	54.5%	Pred. 0.44e-94		
Matches 242	Conservative 82	Mismatches 108	Indels 12	Gaps 4
QY	3	SISMMSGPSYLSLSETLPCNGINGCIKQARKYTVYIGSGDPFASKSLTIRLRGCHVYVIGSN	62	
Db	21	SLHLVSDSDSLAK--VPD-----BAFK-VGILSGGFARSLATRLVSGGFYVVGSHN	70	
QY	63	KPKASEFPFHVVDVTHHEDALTKTVIFVAIAREHYTSLMDLRHLVGKLLIDVSN--NM	120	
Db	71	PKRTALFBSAOVTTQEEAVSSPEVIVAFREHYSLGSLSDQLAGKLLVDVSNPTQ	130	
QY	121	RINQVPSNAEYLASLFPDLSLVKGFNVASAWALQGFKASRGVYICSNNICARQVTE	180	
Db	131	EHQHEESNAEYLASLFPCTVKKAFNVISAMTLQAGRDNRQVPLCGQPKAKVASE	190	
QY	181	LARQMLPIPLDGLSGSAREINLRLFTLLMRGVVAISLAFEFYLSFVADVHPA	240	
Db	191	MALAMGMPDMDGSLASAEVAMPRLPLPAKRPDTLLALDFVCFAPVADVLOFPV	250	
QY	241	RNOQSDFYKPIELIVNKLPIVALITLISLVLAGLAAAYOLYGYKRRFPFMLETWIQ	300	
Db	251	QESQNKPEFLPVGVNVTLLPCVAVYLLSIVLPGLAALQLRGRTKQGFPMLEHMIQ	310	
QY	301	CRQQLGLSFFPMVAVASLCLPRRSERYIFLMAQQVHANISNNEEVMREIVY	360	
Db	311	HRQQLGLSFFCALHALYSCFLPRRARVDLVLAQVLAQKSHLWEEVMEVREIY	370	
QY	361	ISFGMISGLSLSLAATSIPIVSNALNMNREFSIQSTLGVALLISFPHVLIYGMKRAE	420	
Db	371	LSLGVIALGTLISLAAVTSLPSTANSNMNREFSVQSLGFVALVLSLHTLTLYGWMRAE	430	
QY	421	EERYRPTPPNFVALVLPSTIYL 444		
Db	431	ESRYKFLPEPTFTLLVLPQVIL 454		
RESULT 11				
Q9NVB5		PRELIMINARY;	PRT;	488 AA.
ID	Q9NVB5			
AC	Q9NVB5			
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Hypothetical protein FLJ10829 (Dudulin 2).			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catartini; Homiidae; Homo.			
CC	NCBI_Taxid:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Isegal T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,			
RA	Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,			
RA	Nakamura Y., Nagahari K., Maeno Y., Sasaki N.			
RT	"NAKDO human cDNA sequencing project."			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Serru V., Manivet P., Lenoir C., Bescheve P., Lambin D.,			
RA	Vaubourdoille M., Kellermann O., Loric S.;			
RT	"Dudulin 2, a new tumor antigen expressed in various human tumors."			
DR	EMBL; AK001691; BA91839.1; -			
DR	EMBL; AY029585; AAK5038.1; -			
DR	GO; GO:0006118; P:electron transport; IEA.			
DR	InterPro; IPR003006; IG_MHC.			
DR	InterPro; IPR004455; NADPoxred_F420.			
DR	Pfam; PF01807; F420_oxidored_1.			
DR	PROSITE; PS00290; IG_MHC. 1.			
DR	Hypothetical protein.			
SC	SEQUENCE 488 AA; 54616 MW; BC0BCA483335AAD6 CRC64;			

```

Query Match      54.0%; Score 1269; DB 4; Length 488;
Best Local Similarity 54.3%; Pred. No. 7.5e-94;
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;

QY 3 SISMMGSKSLSEFCLENGINGIKDKARKTVGVYISGDFAKSLTIRLRICGHVIVISRN 62
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 11 SLHVLSDSSSLAK--VDP-----EAK--VGISGDFAKSLATRLVSGGFKVYVSRN 60
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 63 PKFASEFPFHVADVYTHHEDALTITKINIFVAIHREHTSLMDRLHVLGKILDVSN--NM 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 PKRTARLVPSAAQYTPQEBAVSSPEVIFVAVPREHYSLSLCSIDQAGKILVDVSNPTQ 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 121 RINQYPSGNAEYLAIFPPDSIYKGFNVVAAALQIGPPDASROYVIGSNNIQAAQYIE 180
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 EHLQRESNAYKLSPETCTVVAFAVVISAMTQAGPPDGKRVPICADQPEAKAAVSE 180
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 181 LARQNFIPIDIGSLSSAREIENLEPLFLMLRGPRVVAISLATEFFLYSFVDYIHRYA 240
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 181 MLAMGFPVDMGSLASMEVEAMPRLRLPAMKVPFLMLLGLFVCYVAFNFRDVLQPYV 240
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 241 RNQGSDFPKIPLEIVNKLTPVIAITLLSLVYLAGLLAAAYQLYGKTKRRPPEMLTQ 300
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 241 QESQKKPFKLPVSVVNTLPCVAAYVLLSLVYLVGLVLAALQHRGTKYQRFEDMLDHWLQ 300
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 301 CRRQGLISFPFAWVAVSLCEPMRSEXYLFLNMAVQOVHANTENSNBEERVIMY 360
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 301 HKRQGLISFPKALHLYSCFLPRAHRYDLVNLAVQVLANKSHLWVEEVARMLY 360
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 361 ISFGIMSLGLSLIAVTSIPSVSNALNWRFFSTQSTLGYALLISTFHVLLYGMKAF 420
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 361 LSLGLVLAIGTSLIAVTSIPSIANSLNMRFFSFVSSLGVAIVLSTHTLLTYGWTAF 420
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 421 EEXYRFPYPPNVVLAIVLPSVIL 444
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 421 ESRKYFVLPPTFTLLVPCVIL 444
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 12
QSNEM6 ID PRELIMINARY: PRT; 487 AA.
QSNEM6 ID QSNEM6 PRELIMINARY: PRT; 487 AA.
AC QSNEM6 PRELIMINARY: PRT; 487 AA.
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE phyde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Y., Beheshati B., Squire J.A., Yang X.J.;
RT "Characterization of a novel apoptosis-inducing gene, hphYde, that
RT inhibits prostate cancer cell growth."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY082673; RAM45136.1; -
DR GO: GO:006118; Pteleictron transport; IEA.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR004455; NADPoxred_F420.
DR Pfam: PF03807; F420_oxidored; 1.
DR PROSITE: PS00290; IG_MHC; 1.
DR SEQUENCE 487 AA; 54501 MW; E0EP95E85C81EDF CRC64;
SQ
Query Match      53.3%; Score 1252.5; DB 4; Length 487;
Best Local Similarity 54.1%; Pred. No. 1.6e-92;
Matches 240; Conservative 83; Mismatches 108; Indels 13; Gaps 5

QY 3 SISMMGSPSLSEFCLENGINGIKDKARKTVGVYISGDFAKSLTIRLRICGHVIVISGN 62
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 11 SLHVLSDSSSLAK--VDP-----EAPKVS--ILGSGDFARSLATRLVSGGFKVYVSGSN 60
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 63 PKFASEFPFHVADVYTHHEDALTITKINIFVAIHREHTYSLMDRLHVLGKILDVSN--NM 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 PKRTARLVPSAAQYTPQEBAVSSPEVIFVAVPREHYSLSLCSIDQAGKILVDVSNPTQ 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 121 RINQYPSGNAEYLAIFPPDSIYKGFNVVAAALQIGPPDASROYVIGSNNIQAAQYIE 180
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 EHLQRESNAYKLSPETCTVVAFAVVISAMTQAGPPDGKRVPICADQPEAKAAVSE 180
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 181 LARQNFIPIDIGSLSSAREIENLEPLFLMLRGPRVVAISLATEFFLYSFVDYIHRYA 240
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 181 MLAMGFPVDMGSLASMEVEAMPRLRLPAMKVPFLMLLGLFVCYVAFNFRDVLQPYV 240
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 241 RNQGSDFPKIPLEIVNKLTPVIAITLLSLVYLAGLLAAAYQLYGKTKRRPPEMLTQ 300
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 241 QESQKKPFKLPVSVVNTLPCVAAYVLLSLVYLVGLVLAALQHRGTKYQRFEDMLDHWLQ 300
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 301 CRRQGLISFPFAWVAVSLCEPMRSEXYLFLNMAVQOVHANTENSNBEERVIMY 360
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 301 HKRQGLISFPKALHLYSCFLPRAHRYDLVNLAVQVLANKSHLWVEEVARMLY 360
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 361 ISFGIMSLGLSLIAVTSIPSVSNALNWRFFSTQSTLGYALLISTFHVLLYGMKAF 420
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 361 LSLGLVLAIGTSLIAVTSIPSIANSLNMRFFSFVSSLGVAIVLSTHTLLTYGWTAF 420
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 421 EEXYRFPYPPNVVLAIVLPSVIL 444
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 421 ESRKYFVLPPTFTLLVPCVIL 444
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

Db      61 PKRTALFFPSAAQVTFQBEAVSSPEVIFVAVFKEHYSLSLSDQLAGKILVDVSNPTQ 120
Qy      121 RINQYBESNAEYLIALFPDLSILVKGFNVSAMALQIGPKDASROYIICSNNIQARQVIE 180
Db      121 EHLQHRESNAEYLIALFPCTCTVKAFNVISAMTLQAGPRDGNQVPICGDPEAKRAVSE 180
Qy      181 LAQOLNFIPIIDIGLSASAREIENLPLRLFTLMRGVVAIALATFFPLISFVADVHHPA 240
Db      181 MALAMGFPMVDMSGSLASAMEVAMPRLPLPAKVPFTLLALGLVCFYANFVADVLPQPV 240
Qy      241 RNOQSDFKPIPIEIVNKTLPVIAITLISLVYLAGLLAAAYQLYGTYKRRFPMELETWQ 300
Db      241 QESQNKFFKLPVSVVNTLPCVAYVLLSVLPGVLAALQLRGTYQRFPMDLHMLO 300
Qy      301 CRKQGLISFFPAMVAVAYSLCLPMRSEERYFLNMAVQOVHANIEANSWNEEVRIMY 360
Db      301 HRKQIGLISFFCALHALYSCFLPLRAHRYDVLNLAQVLANKSHLM-VEEVRMEIY 359
Qy      361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFIOSTLGYVALLSTFHYLYGWRAPF 420
Db      360 LSLGVALGTLISLAVTSLPSIANSLNWRSEFVQSSLGVALVSTLHTLYGWRAPF 419
Qy      421 EERYRFTPPNFVLTALVPSIVIL 444
Db      420 ESRYKFYLPPTFTLLVPCVIL 443

```

RESULT 13

```

Q8TF03 08TF03 PRELIMINARY; PRT; 487 AA.
AC 08TF03;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Tumor suppressor pHyde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Steiner M.S., Alley J.A., Wang C.;
RT "A novel prostate-derived tumor suppressor.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238864; AAL78206.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 487 AA; 54402 MW; AEF16053590E6F68 CRC64;

```

Query March 53.0%; Score 1245.5; DB 4; Length 487;
 Best Local Similarity 53.8%; Pred. No. 5.8e-92;
 Matches 239; Conservative 83; Mismatches 109; Indels 13; Gaps 5;

```

Qy      3 SISMGSPKSLSETCPNGINGIKDARKVTVGVSGDPFAKSLTIRLIRGCVHVVIGSRN 62
Db      11 SLHLVDSDSLAK--VPD-----EAPKVS--ILGSGDPARSLATRLVSGGFKVVGSRN 60
Qy      63 PKFASEFPFHVVDVTHHEDALTKNIIFFVAIHREHYTSLMDRLHLVGLKILIDVSN--NM 120
Db      61 PKRTALFFPSAAQVTFQBEAVSSPEVIFVAVFKEHYSLSLSDQLAGKILVDVSNPTQ 120
Qy      121 RINQYBESNAEYLIALFPDLSILVKGFNVSAMALQIGPKDASROYIICSNNIQARQVIE 180
Db      121 EHLQHRESNAEYLIALFPCTCTVKAFNVISAMTLQAGPRDGNQVPICGDPEAKRAVSE 180
Qy      181 LAQOLNFIPIIDIGLSASAREIENLPLRLFTLMRGVVAIALATFFPLISFVADVHHPA 240
Db      181 MALAMGFPMVDMSGSLASAMEVAMPRLPLPAKVPFTLLALGLVCFYANFVADVLPQPV 240

```

```

Qy      241 RNOQSDFKPIPIEIVNKTLPVIAITLISLVYLAGLLAAAYQLYGTYKRRFPMELETWQ 300
Db      241 QESQNKFFKLPVSVVNTLPCVAYVLLSVLPGVLAALQLRGTYQRFPMDLHMLO 300
Qy      301 CRKQGLISFFPAMVAVAYSLCLPMRSEERYFLNMAVQOVHANIEANSWNEEVRIMY 360
Db      301 HRKQIGLISFFCALHALYSCFLPLRAHRYDVLNLAQVLANKSHLM-VEEVRMEIY 359
Qy      361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFIOSTLGYVALLSTFHYLYGWRAPF 420
Db      360 LSLGVALGTLISLAVTSLPSIANSLNWRSEFVQSSLGVALVSTLHTLYGWRAPF 419
Qy      421 EERYRFTPPNFVLTALVPSIVIL 444
Db      420 ESRYKFYLPPTFTLLVPCVIL 443

```

RESULT 14

```

Q8TDP3 08TDP3 PRELIMINARY; PRT; 456 AA.
AC 08TDP3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE pHyde II.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang C., Alley J.A., Steiner M.S.;
RT "Second human member of pHyde family, Human pHyde II.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF62322; AAM08128.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420_oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 456 AA; 50430 MW; CSE7C7008D52521E CRC64;

```

Query March 45.7%; Score 1074.5; DB 4; Length 456;
 Best Local Similarity 52.9%; Pred. No. 3.2e-78;
 Matches 209; Conservative 75; Mismatches 98; Indels 13; Gaps 5;

```

Qy      3 SISMGSPKSLSETCPNGINGIKDARKVTVGVSGDPFAKSLTIRLIRGCVHVVIGSRN 62
Db      11 SLHLVDSDSLAK--VPD-----EAPKVS--ILGSGDPARSLATRLVSGGFKVVGSRN 60
Qy      63 PKFASEFPFHVVDVTHHEDALTKNIIFFVAIHREHYTSLMDRLHLVGLKILIDVSN--NM 120
Db      61 PKRTALFFPSAAQVTFQBEAVSSPEVIFVAVFKEHYSLSLSDQLAGKILVDVSNPTQ 120
Qy      121 RINQYBESNAEYLIALFPDLSILVKGFNVSAMALQIGPKDASROYIICSNNIQARQVIE 180
Db      121 EHLQHRESNAEYLIALFPCTCTVKAFNVISAMTLQAGPRDGNQVPICGDPEAKRAVSE 180
Qy      181 LAQOLNFIPIIDIGLSASAREIENLPLRLFTLMRGVVAIALATFFPLISFVADVHHPA 240
Db      181 MALAMGFPMVDMSGSLASAMEVAMPRLPLPAKVPFTLLALGLVCFYANFVADVLPQPV 240
Qy      241 RNOQSDFKPIPIEIVNKTLPVIAITLISLVYLAGLLAAAYQLYGTYKRRFPMELETWQ 300
Db      241 QESQNKFFKLPVSVVNTLPCVAYVLLSVLPGVLAALQLRGTYQRFPMDLHMLO 300
Qy      301 CRKQGLISFFPAMVAVAYSLCLPMRSEERYFLNMAVQOVHANIEANSWNEEVRIMY 360
Db      301 HRKQIGLISFFCALHALYSCFLPLRAHRYDVLNLAQVLANKSHLM-VEEVRMEIY 359
Qy      361 ISFGIMSLGLSLAVTSIPSVSNALNWRSEFIO 395
Db      360 LSLGVALGTLISLAVTSLPSIANSLNWRSEFVQ 394

```


CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF186249; AAF17479.1; -
 DR EMBL: AF005053; AAF19150.1; ALT_INT.
 DR EMBL: AC004969; AAD15620.2; -
 DR EMBL: BC011802; AAH11802.1; -
 DR Genew: HGNC:11378; STEAP.
 DR MIM: 604415; -
 DR GO: GO:0005887; C: integral to plasma membrane; TAS.
 DR GO: GO:0005911; C: intercellular junction; TAS.
 DR GO: GO:0015267; F: channel/pore class transporter activity; TAS.
 KM Transmembrane, Antigen.
 FT TRANSMEM 71
 FT TRANSMEM 119
 FT TRANSMEM 164
 FT TRANSMEM 218
 FT TRANSMEM 258
 FT TRANSMEM 291
 FT SEQUENCE 339 AA; 39651 MW; 55443A170C870387 CRC64;

Query Match 30.5%; Score 717; DB 1; Length 339;
 Best Local Similarity 54.9%; Pred. No. 6e-48;
 Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLMRGVVAVALAFPPFYSPFVDVHPYAHNQSDDYKPIPIELVNKPLPIVATTL 267
 DB 67 LFPQWHLPIKAIATIASLTFPIYTLRETHPLATSHOQYFKPIPIVINKVLPWVSTTL 126
 QY 268 SLVYIAGLAAAYOLYVGYTKRPPWLETLQCRKQGLISFFPAAVHVAVSLCPMR 327
 DB 127 ALVLPGLVIAIVQLHNGTKYKPKPWLDKWMLTRKQFGLISFPFAVLHATYSLSYMR 186
 QY 328 SERLFLPAAVQVHANENSNNEEYVRIEMTISFGIMSLGLSLAAVTSISYSNAN 387
 DB 187 SYRKLLNMAVQVQVQNNEDAMIEDHWMEIYVSLDVGAIITALLAVTSISVSLSLT 246
 QY 388 WRESEFIQSTGLVALLISTFPHVLVYGWKRAFEERYFTPTPFLVALVPSIVIL 444
 DB 247 WREHYIYSKGIYSLLGLTHALIPMNKMDIKQVWVTPPFLVAVLPVIVL 303

RESULT 2
 FARE_METUA STANDARD; PRT; 223 AA.
 ID FARE_METUA
 AC Q58896;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Putative F420-dependent NADP reductase (EC 1.-.-.-).
 GN MJ1501.
 OS Methanococcus jannaschii.
 CC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 CC Methanocaldococcaceae; Methanocaldococcus.
 CC NCBI_TaxID=2190;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RC MEDLINE=96337993; PubMed=8688087;
 RX Buit C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kesteven A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Fuhmann J.L., Nguyen D.,
 RA Scott J.L., Georgiades N.S.M., Weidman U.F., Fuhrmann J.L., Nguyen D.,
 RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Coton M.D., Roberts K.M., Hurt M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-G., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii";
 RT Science 273:1058-1073 (1996).
 RL CC -I- FUNCTION: Catalyzes the reduction of F420 with NADP(+) and the

CC reduction of NADP(+) with F420H(2). Probably couples the NADP-
 CC dependent oxidation of the alcohol to the aldehyde function (By
 CC dependent reduction of CO(2) to methane (anaerobic function) (By
 CC similarity).
 CC -I- CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced
 CC coenzyme F420.
 CC -I- SIMILARITY: TO M.ORGANOPHILUM F420-DEPENDENT NADP REDUCTASE.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U67591; AAB99514.1; -
 DR PIR: D64487; D64487.
 DR TIGR: M1501; -
 DR InterPro: IPR004455; NADPoxred_F420.
 DR Pfam: PR03807; F420 oxidored; 1.
 DR TIGRFAMs: TIGR00301; TIGR00301; 1.
 KM Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
 DR SEQUENCE 223 AA; 24068 MW; 2370BD6F5BD3D9 CRC64;

Query Match 5.7%; Score 133; DB 1; Length 223;
 Best Local Similarity 21.9%; Pred. No. 0.003;
 Matches 46; Conservative 52; Mismatches 78; Indels 34; Gaps 7;

QY 37 GSGDPAKSTIRLNGGVHVGSGNPKFASFPFHVDV-----THREDALT 84
 DB 8 GTDGDGFGIARLAK-NKKIIGSKRKEKAEPAKAKVILKQGIADITGLEKDAK 66
 QY 85 KNIIFVAIHREH-YTSLMDIRHLVGLIDV-----SNMRINQYE-SNMEYL 133
 DB 67 EGDVILSLPYVYITSLTKQLKELKGIIVSIGVPLATAGDKPTLFPDDSVAMV 126
 QY 134 ASLFPDLSLVKGFNVYSMAALQKPKDASROYIICSNNOQROYIELARLNT-PIDL 192
 DB 127 QNVLKESKVSFAFQVYCHAVVEDIDNPVDCDILVGDNEAKVYIDLANIDGVRALDC 186
 QY 193 GSLSSARETENDPLRLFTLMRGVVAISL 222
 DB 187 GNLEKSRITTEALT-----PLIGLNT 207

RESULT 3
 FARE_MERTH STANDARD; PRT; 232 AA.
 ID FARE_MERTH
 AC Q26350;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Putative F420-dependent NADP reductase (EC 1.-.-.-).
 GN MTH248.
 OS Methanobacterium thermoautotrophicum.
 CC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 CC Methanobacteriaceae; Methanobacter.
 CC NCBI_TaxID=187420;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Delta H;
 RC MEDLINE=98037514; PubMed=9371463;
 RX Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lunn W., Pothier B., Qiu D.,
 RA Spadefora K., Vickre K., Wang Y., Wierzbowski J., Gibson R.,
 RA Dwan N., Caruso A., Bush D., Sifer H., Patwell D., Fradnakar S.,
 RA McDougall S., Shiner G., Goyal A., Pletrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RA "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT delah; functional analysis and comparative genomics";
 RT J. Bacteriol. 179:7135-7155 (1997).
 RL CC

```

CC -1- FUNCTION: Catalyzes the reduction of F420 with NADP(+) and the
CC reduction of NADP(+) with F420H(2). Probably couples the NADP-
CC dependent oxidation of the alcohol to the aldehyde with the F420-
CC dependent reduction of CO(2) to methane (anabolic function) (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced
CC coenzyme F420.
CC -1- SIMILARITY: TO M.ORGANOPHILUM F420-DEPENDENT NADP REDUCTASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000811; AAB84754.1; -.
CC PIR: A69131; A69131.
CC InterPro: IPR004455; NADPoxred_F420.
CC Pfam: PF03807; F420_oxidored_1.
CC TIGRFAMs: TIGR00301; TIGR00301.1.
CC DR TIGRFAMs: TIGR00301; TIGR00301.1.
CC KW Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
SQ SEQUENCE 232 AA; 24539 MW; A1CE60ABC8474296 CRC64;

Query Match 5.1%; Score 120.5; DB 1; Length 232;
Best Local Similarity 23.8%; Pred. No. 0.029; 90; Indels 25; Gaps 6;
Matches 49; Conservative 42; Mismatches 90;

QY 30 KVTGVIGSGDPFANSLTIRLCGYHYVVGSRNPKFSEPPHYVDVTHHE----- 80
DB 8 KIAV-IGTGDOGGLAFPAVAGEVITIGSDAEKASKAYLEAGRDISEGATN 66
QY 81 -DALTKNIIIVAIH-REHYTSLMDRLHLVGLKILIDS-----NMNRINQYPSN 129
DB 67 PDAAASAVVVVLTPLQAVYTLASIRQVDRKVLIDFTVDSICIGSAVRYYDMEGS 126
QY 130 AEYLASLF--PDSLIVGPFVVSAMALQPKDASRQVYICSNNOIQRCQVIELARLN 186
DB 127 AAEARARLRQGTGTVAAAFNNISASALLESEPDCCVLAASHRDALVPAALAEKID 186
QY 187 FI-PIDLGSLASAREIENLPLRLFTL 211
DB 187 GVRALCGGLNARIIEKITPLINL 212

RESULT 4
FSHR_MACFA STANDARD; PRT; 695 AA.
ID_FSHR_MACFA STANDARD; PRT; 695 AA.
AC P32212;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictrophin
DE receptor).
GN FSHR.
OS Macaca fascicularis (Crest eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OC NCBI_TaxID=9541;
CX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Testis;
RX MEDLINE=94071854; PubMed=7504463;
RA "Gronoll J., Dankar B., Sharma R.S., Nieschlag E.;
RT "Molecular cloning of the testicular follicle stimulating hormone
RT receptor of the non human primate Macaca fascicularis and
RT identification of multiple transcripts in the testis.";
RL Biochem. Biophys. Res. Commun. 196:1066-1072(1993).
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
CC of this receptor is mediated by G proteins which activate
CC adenylyate cyclase.

```

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC FSH/LSH/TSH subfamily.
CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X74454; CAA52463.1; -.
CC PIR: UN0898; UN0898.
CC HSSP: F23945; LXUN.
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR000372; LRR_Nterm.
CC Pfam: PF00001; 7tm.1.1.
CC Pfam: PF00560; LRR_3.
CC Pfam: PF01462; LRRNT.1.
CC DR PRINTS; P000237; GPCRHDOPSN.
CC DR SMART; SMO0013; LRRNT.1.
CC DR PROSITE; PS00237; G PROTEIN RECP_F1_1; 1.
CC DR PROSITE; PS02622; G PROTEIN RECP_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1..17
FT CHAIN 18..695
FT DOMAIN 18..366
FT TRANSMEM 367..387
FT DOMAIN 388..398
FT TRANSMEM 399..421
FT DOMAIN 422..443
FT TRANSMEM 444..465
FT DOMAIN 466..485
FT TRANSMEM 486..508
FT DOMAIN 509..528
FT TRANSMEM 529..550
FT DOMAIN 551..573
FT TRANSMEM 574..597
FT DOMAIN 598..608
FT TRANSMEM 609..630
FT DOMAIN 631..695
FT REPEAT 44..68
FT REPEAT 69..93
FT REPEAT 119..143
FT REPEAT 170..192
FT REPEAT 193..216
FT REPEAT 218..240
FT DISULFID 442..517
FT CARBOHYD 191..191
FT CARBOHYD 199..199
FT CARBOHYD 293..293
FT CARBOHYD 318..318
SQ SEQUENCE 695 AA; 78343 MW; 0D60A233729B5250 CRC64;

Query Match 4.7%; Score 111.5; DB 1; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.56;
Matches 84; Conservative 75; Mismatches 151; Indels 183; Gaps 16;

QY 79 HEDALTKNIIIVAIHREHYTSLMDRLHLVLC-----KILIDVSNMR 121
DB 98 HETRIKANNL-LYNPEAFQNLPLRYLLISNTGIKELPDVHKHSFQKVLDDIQDNTN 156
QY 122 INQYPSNAEYASLPDLSIVGFVVSAMALQPKDASRQVYICSNNOIQRCQVIEL 181
DB 157 IH-----TIRNSFVGSLFESVILW-----NKGIGIEINCA----- 189
QY 182 ARQNLTPIDLGSLASAREIENLPLRLFTLMGPVVAISLTFPVSFVADVHPYAR 241
DB 190 ---FNGTQIDNLSDNNNLIELPNDVFGASGPVILDISRTIRHSLPSYGLNKKJRA 246

```

```

QY 242 NQGSDFYKIPLEIVNKTLPVIAITLSLVY----- 271
DB 247 RSTYNLKLP-----SLEKVALMEASITPSHCCAFANWRROISELHPICNKSILRQEV 301
QY 272 -----LAGLLAAAYQVLYGKRYR-----PEMLTWTQCRKQLG 306
DB 302 DYMOTGTGQSSLDNEDSSYSRGDMTYAEFDYDLCNEVVDVTCSPKPAFNPCEIDILG 361
QY 307 -----LISFFPAMVHA-----YSLCPFRSERRYFLNMAVQVHANIENS 348
DB 362 YNLRVLIWFISLAIITNIIIVLTLTSQYKLTVP-----RFLMCLAFAD----- 408
QY 349 WNESEWRIEMYISFGIMSLGLSLAVTSIPSVSNALNRE-----FSTIOSTLG 339
DB 409 -----LCIGYLLILIASVDIHTKSQYHNVAIDWGTAGCDPAGCFVFASELS 456
QY 400 -----YVALISTFPH-----VLIGMKAFEEERYRFPNPFV-LAL 436
DB 457 VYTLTAITLERHHTITHAMQLDCKVHVHAAVWVMGMIFAPALPFIIGISSYMKVSI 516
QY 437 VLPSVILDLIQL 449
DB 517 CLPMDIDSPLSQL 529

```

RESULT 5

```

ID Y538_PASMTU STANDARD; PRT; 206 AA.
AC Q9CN97;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein PM0538.
GN PM0538.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genome sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -|- SIMILARITY: Belongs to the UPF0191 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; AB006089; AA002622.1; -
CC HAMAP; MF_01207; -; 1.
CC InterPro; IPR007916; UPF0191.
CC Pfam; PF05252; UPF0191; 1.
CC
CC Hypothetical protein; Transmembrane; Complete proteome.
CC
CC TRANSMEM 7 26
CC TRANSMEM 41 63
CC TRANSMEM 76 98
CC TRANSMEM 113 135
CC TRANSMEM 142 164
CC TRANSMEM 169 191
CC
CC SEQUENCE 206 AA; 23653 MW; 2F277729D7F11194 CRC64;

```

Query Match 4.7%; Score 109.5; DB 1; Length 206;
 Best Local Similarity 23.8%; Pred.No.0.18; 72; Indels 63; Gaps 13;
 Matches 56; Conservative 44; Mismatches 63; Gaps 13;

```

QY 228 LYSFVRDVIH-----PYA-----RNOQSDPYKIPLEIVNKTLPVIAITLSLVYLAGL 275
DB 1 MSLSEIRIITHWCCLOPVMAMVLLSGESQLGADPIEIGHFQFSALTILLIMFLIG- 59
QY 276 LAAAYQVLYGTYKRPFPMLLETWLCQRQLGLSFFPAMVHAASCLPMRSERYFLN 335
DB 60 -----KVFLTK-----QPOLQV-----LRPALGLMANFVVLHV-YA-----YALE 96
QY 336 MAYQVHANIENSWNESEWRIEMYISFGIMSLGLSLAVTSIPSVSNALNREESTIQ 395
DB 97 LGYD-----FSLFQELVNR--GYLIIGALAFILITLMAISWSYLR--LKMGMWYEL 146
QY 396 STLGYVALLISTFPHLIGMKRAFEERYRFPNPFVLAIVSLVTLDLQLC 450
DB 147 HQLGYYALLLGAH--YVM-----SVKNVTFSSML-YLILSTMLIC 184

```

RESULT 6

```

ID FSHR_CHICK STANDARD; PRT; 693 AA.
AC P79763; Q90719;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
DE receptor).
GN FSHR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=97473503; PubMed=9332357;
RA Wakabayashi N., Suzuki A., Hoshino H., Nishimori K., Mizuno S.;
RT "The cDNA cloning and transient expression of a chicken gene encoding
RT a follicle-stimulating hormone receptor.";
RL Gene 197:121-127(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97057887; PubMed=8902217;
RA You S., Bridgman J.T., Foster D.N., Johnson A.L.;
RT "Characterization of the chicken follicle-stimulating hormone
RT receptor (FSH-R) complementary deoxyribonucleic acid, and expression
of cFSH-R messenger ribonucleic acid in the ovary.";
RL Biol. Reprod. 55:1055-1062(1996).
CC -|- FUNCTION: Receptor for follicle stimulating hormone. The activity
CC of this receptor is mediated by G proteins which activate
CC adenylate cyclase.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC FSH/LSH/TSH subfamily.
CC -|- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; D87871; BA13487.1; -
CC EMBL; U51097; AAC60030.1; -
CC HSSP; P23945; 1XUN.
CC InterPro; IPR000276; GPCR_Rhodopn.
CC InterPro; IPR001611; LRR_Nterm.
CC Pfam; PF00001; 7tm_1; 1.

```

DR Pfam: PF00560; LRR; 2.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR SMART: SMO0013; LRRNT: 1.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR; 1.
 DR PROSITE: PS50262; G-PROTEIN RECEPTOR; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
 KW phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 17
 FT CHAIN 18 693
 FT DOMAIN 18 366
 FT TRANSSEM 367 387
 FT DOMAIN 388 398
 FT TRANSSEM 399 421
 FT DOMAIN 422 443
 FT TRANSSEM 444 465
 FT DOMAIN 466 485
 FT TRANSSEM 486 508
 FT DOMAIN 509 528
 FT TRANSSEM 529 550
 FT DOMAIN 551 573
 FT TRANSSEM 574 597
 FT DOMAIN 598 608
 FT TRANSSEM 609 630
 FT DOMAIN 631 663
 FT REPEAT 45 98
 FT REPEAT 69 93
 FT REPEAT 95 118
 FT REPEAT 119 143
 FT REPEAT 168 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 47 191
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 268 268
 FT CONFLICT 4 4
 FT CONFLICT 88 88
 FT CONFLICT 140 140
 FT CONFLICT 174 174
 FT CONFLICT 191 191
 FT CONFLICT 329 329
 FT SEQUENCE 693 AA; 78697 MW; 46F9869635A1BEC64;
 Query March 4.6%; Score 109; DB 1; Length 693;
 Best Local Similarity 17.2%; Pred. No. 0.87; Indels 238; Gaps 21;
 Matches 100; Conservative 72; Mismatches 173; PKL-----PKL----- 97
 QY 18 LPNG-INGIKDARKVTG-----VIGSGDPAKSLTIRILRCGVHVIIGSRNPKPASEFFP 71
 DB 61 IPKGFETLDHLEKIEIQNDALIEIGNVPSL-----PKL----- 97
 QY 72 HAVDVTHEDALTKNITFVAIHEHTSLMDLHLVVG-----KILI 114
 DB 98 -----HEIRIEKXNKU-MKIDDAFQHLPSRLYLSNTGSLFPLVPEKHSFOKVL 149
 QY 115 DVSNMRINQYPSNAEYIASLFPDSLIVKGFNVSAWALQGPKDSRQYIYCSNNIOA 174
 DB 150 DVQDNHRTIEHNFMLSS-----ESVILR-----LNKGIOE 184
 QY 175 RQQVIELRQNFITIDGSLSSAEIENLRLFTLMRGVVAIASLATFFLYSEVRD 234
 DB 185 IKD-----HAFNGTGLDELINSDNYNLEKPEKVFQAGVAVDISTRISFLPSHGLE 239
 QY 235 VTHPYARNOQSDFYKIPLEIYNKTLPIVAITLLSVIYAG----- 275
 DB 240 FIKKURASTYKTKLP--DVNKKFSLSEANFTYSHCCATNKTQNTERTYPICSMSPA 297
 QY 276 -----LAAVQLYGYTK-----RFF-----PWLFTWLOCR 302
 DB 298 KQDLGEQTKRKRGRSAEDYISHYGTBFGPVENEFYDGLCNEVVDVFCSEKPPAFNCE 357
 QY 303 KQLG-----LISFFAMVHA-----YSLCLPMRSEBRYLFNNMAVQOVHAN 344

DB 358 DIMGVNVLRLVILWFIFINIALIGNTVLLIILISSQKVLVP-----RFIMCMLAFAD 408
 QY 345 IINSNNEDEWRIENYIFGIMSLGLSLAVTSIPSUSNLANRE-----FSFIQ 395
 DB 409 -----LCIGILYLFASVDIQTSRYNNYVADMQTGACGNAAGFFYFA 452
 QY 396 STLGVALLISTFH-----VLIYMKCAF----- 419
 DB 453 SELSVYTLVTLERHMTITYAMQLNRKRYLRHAYIIVFGMPFETALLPIFGISSYM 512
 QY 420 -----EEHYRYFTPPNFVATVLPISVILDLQCR 452
 DB 513 KVSICLPMHETPFQAVV--IFLIVNLAVIICICICYIC 553
 RESULT 7
 ID FRB6_YEAST STANDARD; PRT; 712 AA.
 AC Q12473;
 DT 15-DEC-1998 (rel. 37, Last sequence update)
 DT 15-DEC-1998 (rel. 37, Last sequence update)
 DT 10-OCT-2003 (rel. 42, Last annotation update)
 DE Ferric reductase transmembrane component 6 precursor (EC 1.16.1.7)
 DE (ferric-chelate reductase 6).
 GN FRB6 OR YIL051C OR I0593.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=9731367; PubMed=9169871;
 RA Johnson M., Hillier L., Riles L., Albertmann K., Andre B., Ansgorge W.,
 RA Benes V., Bruckner M., Deltus H., Dubois E., Duesterhoeft A.,
 RA Eutian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
 RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
 RA Louis E.J., Messenguy F., Mewes H.-W., Mosga T., Moestl D.,
 RA Mueller-Auer S., Newnich U., Obermaier B., Piravandi E., Pohl T.M.,
 RA Portetelle D., Purnelle B., Reckmann S., Rieger M., Rinke M., Rose M.,
 RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
 RA Underwood A.P., Urrutiarazu L.A., Vandembol M., Verhaeselt P.,
 RA Viereckels F., Voet M., Volckaert G., Voss H., Wambolt R., Wedler E.,
 RA Wedler H., Zimmermann F.K., Zollner A., Han J., Hohnselt J.D.;
 RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
 RT Nature 387:87-90(1997).
 RL Nature 387:87-90(1997).
 CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
 CC -1- COFACTOR: FAD (Probable).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: Belongs to the FRE / CYB family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: Z47973; CAA88006.1; -
 DR EMBL: Z73156; CAA97503.1; -
 DR PIR: S50969; S50969.
 DR GERMOnline: 142046; -
 DR SGD: S00003974; FRB6.
 DR InterPro: IPR002916; Ferric_reduct.
 DR Pfam: PF01794; Ferric_reduct. 1.
 KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;
 KW PDB: NAD; Glycoprotein; signal; Multigene family.
 FT SIGNAL 1 17
 FT CHAIN 18 712
 FT NP_BIND 493 499
 FT FAD (POTENTIAL).
 FT FERRIC REDUCTASE TRANSMEMBRANE COMPONENT

FT TRANSMEM 168 188 POTENTIAL.
 FT TRANSMEM 245 265 POTENTIAL.
 FT TRANSMEM 268 308 POTENTIAL.
 FT TRANSMEM 329 349 POTENTIAL.
 FT TRANSMEM 361 381 POTENTIAL.
 FT TRANSMEM 388 408 POTENTIAL.
 FT TRANSMEM 417 437 POTENTIAL.
 FT TRANSMEM 478 498 POTENTIAL.
 FT TRANSMEM 550 570 POTENTIAL.
 FT DOMAIN 553 570 POT-LEU.
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 124 124 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 655 655 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 712 AA; 81989 MM; 5224F12B51544ABA CRC64;

Query Match 4.6%; Score 109; DB 1; Length 712;
 Best Local Similarity 18.8%; Pred. No. 0.9; Indels 142; Gaps 17;
 Matches 72; Conservative 51; Mismatches 119;

QY 88 IIPVAIHREHYTSL-----MDLRHLVGLKILIDVSNMRINQYPSNAEYLA SL 136
 DB 183 IIAVFPFHSHNGNRLAFASPVNYIRGHFVLPTEFLVD---KHAHFPELVNEVFTGL 238
 QY 137 FPDGLIYKGFNVVSNALQLGPKDASROVYICSNINQARQVIELAQNLPIPDLSLS 196
 DB 239 MPNSL-----EAW----- 246
 QY 197 SARIEHNPRLFTLMRGPVVAISLAFPPFL-YSFADVIHPARNOQSPFYKIPLEIV 255
 DB 247 -----IIFGYTLANIIILSLISYIIDPNNILFNSHLSQFTSL--LA 284
 QY 256 NKTLPVIAITLISLVYLAGLAAVQLYGTGKRRFPFWLTMQCKRQGLSFPFAMV 315
 DB 285 DRS-GILAFLOPPLIIFTRANSPLEPLGVKFNFSF-----ISFHWIGRIWLNATI 336
 QY 316 H-VAVSICLPMRSEHYFLNMAVQOVHANENSNNEEVRIMYISFGIMSLGLSL 374
 DB 337 HSLSLSL-----FALINAEK-----ISNK-----GLYKFGIASLTVICVL 373
 QY 375 AVTSPYSNALNWEESFIOSTIGYVALLI--STFVLLIY-GMKCAF-----EE 422
 DB 374 IVLISGIYRK-----RHYEFPLVTHIILALFFYCCWQVKLFGNKWIVVSLIWGLEK 429
 QY 423 YVRFYTPNFVIALVLPISIVIL 446
 DB 430 LFRINW---LIQFRFPKATLNL 449

RESULT 8
 FRET YEAST STANDARD; PRT; 629 AA.

ID FRET YEAST
 AC 012333;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ferric reductase transmembrane component 7 (EC 1.16.1.7) (Ferric-
 DE chelate reductase 7).
 GN FRET OR YOL152M.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NCBI_TaxID=4932;
 [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=8288C / FY1679;
 RX MEDLINE=96132030; PubMed=8553699;
 RA Casanueva A., Aldea M., Casas C., Herrero E., Gamo F.J.,
 RA Lafuente M.J., Garcido C., Arino J.,
 RA "DNA sequence analysis of a 13 kbp fragment of the left arm of yeast
 RT chromosome XV containing seven new open reading frames.";
 RL Yeast 11:1281-1288(1995).

CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
 CC -1- COPACITOR: FAD (Probable).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the FRS / CYBB family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z48239; CAA88276.1; -
 DR EMBL; Z74894; CAA89174.1; -
 DR FIR; S60385; S60385.
 DR Geronline; 143574; -
 DR SCD; S0005512; FRET.
 DR InterPro; IPR002916; Ferric reduct.
 DR Pfam; PF01794; Ferric reduct; 1.
 KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;
 KW FAD; NAD; Glycoprotein; Multigene family;
 KW NE BIND FAD (POTENTIAL).
 FT TRANSMEM 46 66 POTENTIAL.
 FT TRANSMEM 108 128 POTENTIAL.
 FT TRANSMEM 168 188 POTENTIAL.
 FT TRANSMEM 195 215 POTENTIAL.
 FT TRANSMEM 238 258 POTENTIAL.
 FT TRANSMEM 266 286 POTENTIAL.
 FT TRANSMEM 293 313 POTENTIAL.
 FT TRANSMEM 422 442 POTENTIAL.
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 541 541 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 629 AA; 71936 MM; 2384480B9289C16F CRC64;

Query Match 4.6%; Score 107.5; DB 1; Length 629;
 Best Local Similarity 21.7%; Pred. No. 1;
 Matches 57; Conservative 46; Mismatches 77; Indels 83; Gaps 17;

QY 214 GPVVAISLAFPPFLVSPFVDVIHPARNOQSPFYKIPLEIVNKTLPVIAITLISLVY-L 272
 DB 123 GTFLVVMATITVLTLLCYFP--HPFYR-PCAGFSPPLSV--RAGIMALSLVFPVPSL 175
 QY 273 AG-----LAAVQLYGTGKRRFPFWLTMQCKRQGLSFPFAMVAVSLCPM 325
 DB 176 SGKINIVGWLGVSYE-----KINIVHW-----ASTLCFFSVHW-----IF 215
 QY 326 RRSERYLFNMAVQOVHANENSNNEEVRIMYISFGIMSLGLSLAVTSIP----- 380
 DB 216 LRGARH---EGYERH-----QWKASDMR-----SGVPILFLNMLSLPIARRH 262
 QY 381 -SVSNALNWEESFIOSTIGYVALLISFHY-----LIYGRKAFEEYR- 425
 DB 263 FYEIFQLHW-----LVAAGFYISLP---YHYVPELNSHMYLVATIVW---FAQLFRL 311
 QY 426 ----FYTPNFVIALVLPISIVIL 444
 DB 312 AVKGYLRPGRSFMASITAVSV 334

RESULT 9

FSHR HUMAN STANDARD; PRT; 695 AA.

ID FSHR HUMAN
 AC P23945; 016225;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicitropin
 DE receptor).
 GN FSHR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

CX NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS ALA-307 AND SER-680.
 RC TISSUE=Ovary;
 RX MEDLINE=91222171; PubMed=1709010;
 RA Minegishi T., Nakamura K., Takakura Y., Ibuki Y., Igarashi M.;
 RT "Cloning and sequencing of human FSH receptor cDNA.";
 RL Biochem. Biophys. Res. Commun. 175:1125-1130(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93246012; PubMed=1301382;
 RA Kelson C.A., Cheng S.V., Nugent N.P., Schweichardt R.L.,
 RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzaja J.B., Luchette C.A.,
 RA Chappel S.C.;
 RT "The cloning of the human follicle stimulating hormone receptor and
 RT its expression in COS-7, CHO, and Y-1 cells.";
 RL Mol. Cell. Endocrinol. 89:141-151(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tilly L.T., Ahnara T., Nishimori K., Jai X.-C., Billig H.,
 RA Kowalski K.I., Perlas E.A., Heueh A.J.;
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-342 FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93075197; PubMed=1359889;
 RA Gromoll U., Gudemann T., Nieschlag E.;
 RT "Molecular cloning of a truncated isoform of the human follicle
 RT stimulating hormone receptor.";
 RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).
 RN [5]
 RP SEQUENCE OF 1-51 FROM N.A.
 RX MEDLINE=95011044; PubMed=7926278;
 RA Gromoll U., Dankhar B., Gudemann T.;
 RT "Characterization of the 5' flanking region of the human follicle-
 RT stimulating hormone receptor gene.";
 RL Mol. Cell. Endocrinol. 102:93-102(1994).
 RN [6]
 RP SEQUENCE OF 286-695 FROM N.A., AND VARIANT SER-680.
 RX MEDLINE=9500244; PubMed=7916967;
 RA Gromoll U., Ried T., Holtgreve-Grez H., Nieschlag E., Gudemann T.;
 RT "Localisation of the human FSH receptor to chromosome 2p21 using a
 RT genomic probe comprising exon 10.";
 RL J. Mol. Endocrinol. 12:265-271(1994).
 RN [7]
 RP 3D-STRUCTURE MODELING OF 49-228.
 RX MEDLINE=96363672; PubMed=8747461;
 RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H.,
 RA Hendrickson W.A., el Tayar N.;
 RT "Structural predictions for the ligand-binding region of glycoprotein
 RT hormone receptors and the nature of hormone-receptor interactions.";
 RL Structure 3:1341-1353(1995).
 RN [8]
 RP VARIANTS ALA-307, ARG-524 AND SER-680.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes.";
 RL Nat. Genet. 22:231-238(1999).
 RN [9]
 RP ERBATUM.
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.O.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).
 RN [10]
 RP VARIANTS ALA-307 AND SER-680.
 RX MEDLINE=22054685; PubMed=12059813;

RA Aestiant K., Gromoll U., Eckardstein S.V., Zitzmann M., Nieschlag E.,
 RA Simon M.;
 RT "Distribution and function of FSH receptor genetic variants in normal
 RT men.";
 RL Andrologia 34:172-176(2002).
 RN [11]
 RP VARIANT OHSS ILE-449.
 RX MEDLINE=22812036; PubMed=12930927;
 RA Vasseur C., Rodien P., Beau I., Desroches A., Gerard C.,
 RA de Poncheville L., Chaplot S., Savagner F., Croue A., Mathieu E.,
 RA Lahou N., Descamps P., Mizrahi M.;
 RT "A chorionic gonadotropin-sensitive mutation in the
 RT follicle-stimulating hormone receptor as a cause of familial
 RT gestational spontaneous ovarian hyperstimulation syndrome.";
 RL New Engl. J. Med. 349:753-759(2003).
 RN [12]
 RP VARIANT OHSS ASN-567.
 RX MEDLINE=22812037; PubMed=12930928;
 RA Smith G., Olatunbosun O., Delbore A., Pierson R., Vassart G.,
 RA Costagliola S.;
 RT "Ovarian hyperstimulation syndrome due to a mutation in the
 RT follicle-stimulating hormone receptor.";
 RL New Engl. J. Med. 349:760-766(2003).
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
 CC of this receptor is mediated by G proteins which activate
 CC adenylate cyclase.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P23945-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P23945-2; Sequence=VSP_001953, VSP_001954;
 CC -1- TISSUE SPECIFICITY: Sertoli cells and ovarian granulosa cells.
 CC -1- DISEASE: Defects in FSHR are a cause of ovarian hyperstimulation
 CC syndrome (OHSS) [MIM:608115]. OHSS is a disorder which occurs
 CC either spontaneously or most often as an iatrogenic complication
 CC of ovarian stimulation treatments for in vitro fertilization. The
 CC clinical manifestations vary from abdominal distention and
 CC discomfort to potentially life-threatening, massive ovarian
 CC enlargement and capillary leak with fluid sequestration.
 CC Pathologic features of this syndrome include the presence of
 CC multiple serous and hemorrhagic follicular cysts lined by
 CC luteinized cells, a condition called hyperreactio luteinalis.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC FSH/LSH/TSH subfamily.
 CC -1- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M65085; AAA52477.1; -
 CC EMBL: S59900; AAB26480.1; -
 CC EMBL: M95489; AAA52478.1; -
 CC EMBL: X68044; CAA48179.1; -
 CC EMBL: S73199; AAB32071.1; -
 CC EMBL: S73526; AAB32225.1; -
 CC PIR: I57661; ORHUPT
 CC PDB: 1XUN; 15-MAY-97.
 CC Genew: HGNC:3969; FSHR.
 CC MIM: 136435; -
 CC MIM: 608115; -
 CC GO: GO:0016021; C:integral to membrane; TAS.
 CC GO: GO:0005886; C:plasma membrane; TAS.
 CC GO: GO:0004963; F:follicle stimulating hormone receptor activity; TAS.
 CC GO: GO:0007292; P:female gamete generation; TAS.
 CC GO: GO:0005885; P:female gonad development; TAS.
 CC GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; TAS.


```

FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 695 FOLICLE STIMULATING HORMONE RECEPTOR.
FT DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 1 (POTENTIAL).
FT DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 399 421 2 (POTENTIAL).
FT DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 444 465 3 (POTENTIAL).
FT DOMAIN 466 485 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 486 508 4 (POTENTIAL).
FT DOMAIN 509 528 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 529 550 5 (POTENTIAL).
FT DOMAIN 551 573 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 574 597 6 (POTENTIAL).
FT DOMAIN 598 608 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 609 630 7 (POTENTIAL).
FT DOMAIN 631 695 CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 68 LRR 1.
FT REPEAT 69 93 LRR 2.
FT REPEAT 119 143 LRR 3.
FT REPEAT 170 192 LRR 4.
FT REPEAT 193 216 LRR 5.
FT REPEAT 218 240 LRR 6.
FT DISULFID 442 517 BY SIMILARITY.
FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 2 2 S -> A (IN REF. 1).
FT CONFLICT 13 13 T -> S (IN REF. 1).
FT CONFLICT 60 60 V -> A (IN REF. 1).
FT CONFLICT 166 166 V -> M (IN REF. 1).
FT CONFLICT 215 215 Q -> H (IN REF. 1).
FT CONFLICT 247 247 K -> R (IN REF. 1).
FT CONFLICT 257 257 S -> T (IN REF. 1).
FT CONFLICT 334 334 D -> N (IN REF. 1).
FT CONFLICT 349 349 E -> K (IN REF. 1).
FT CONFLICT 352 352 T -> A (IN REF. 1).
FT CONFLICT 363 363 V -> E (IN REF. 1).
FT CONFLICT 407 407 A -> T (IN REF. 1).
FT CONFLICT 421 421 V -> I (IN REF. 1).
FT CONFLICT 427 427 T -> S (IN REF. 1).
FT CONFLICT 435 435 D -> N (IN REF. 1).
FT CONFLICT 483 483 L -> V (IN REF. 1).
FT CONFLICT 550 550 T -> I (IN REF. 1).
FT CONFLICT 550 550 A -> V (IN REF. 1).
FT CONFLICT 586 586 A -> V (IN REF. 1).
FT CONFLICT 607 607 S -> L (IN REF. 1).
FT CONFLICT 691 691 R -> H (IN REF. 1).
SQ SEQUENCE 695 AA; 78172 MW; E9EBED829C79C450 CRC64;

```

```

Query Match 4.4%; Score 104.5; DB 1; Length 695;
Best Local Similarity 16.5%; Pred. No. 1.9; Indels 199; Gaps 21;
Matches 91; Conservative 91; Mismatches 169;

```

```

QY 18 LKNG-INGIKDARKVVGIVGSGDPAKSLTIRLCRGYVVGSGNPKFASFFPHVDV 76
DB 61 IKGAFSGGCDJEK-----EISQNVLEIEN-----VFSNLPKL----- 97
QY 77 THHEDALTYNTIIFVAIHREHYTSMDLHLVG-----KILIDVSN 119
DB 98 --HEIRIEKANNIL-LYIDDAFQNLPLRLYLLISNTGVHLPAVKIOSLQRYVLIDIDN 154
QY 120 MNIQVPEENAEYLAFLPDSLIVKGFNVASAMALQPKDASROYVICSNNICARQOYT 179
DB 155 INTH-----TVERNSTFGSLFSBMIML-----SKNGIREIHNA----- 189
QY 180 ELARQINFTPIDGSLSASAREENLPLRLFTLRGCVVAIALATFFLYSVRVDIHBY 239
DB 190 -----FNGQDLDELNTSDNDNLEELPNVDYFGAGSGVILIDISRTRIHSLPSYGLNKL 244
QY 240 ARNQGSDFYKIP-----IEIVNKT-----PIVAITL----- 267
DB 245 RAKSTYLNKLPLSELEFVTLMENSLYTPSHCCAFAMWRQISDLHPTCKSLIRGEVDVM 304

```

```

QY 268 -----SLVYLAGLAAQVLYGTYKRYF-----PMLFTWLCRKOLG--- 306
DB 305 TQARGRVSLAEGSSSLAKEPDTMYSEFDYDLCNEVDVICSPEPDITNPEDIMGHI 364
QY 307 --LISFFPAMVHA-----YSLCLPMRSEERYFLNAYQOVHANISWNE 351
DB 365 LRLVIFILSLATIGNIIVILITSQYKLTVP-----RFLMGNLAFAD----- 408
QY 352 EEVWRLEMTISQIMSLGLSLAATSIPSVNALNWR-----SFISQTLG--- 399
DB 409 -----LCIGTYLLTASVDIHKTQYHNVAIDWQAGCDAAAGFTVASELSYVT 459
QY 400 YVALISTFH-----VLIVGWRAPAEERYRYPDPNV-LALVLP 439
DB 460 LFTALTEPRHTITHAMQLOCKVQLRRHASIMLVGNIFATVALPIFGISSWVKVSI 519
QY 440 STYILDLSQI 449
DB 520 MDIDSPLSQL 529

```

```

RESULT 11
SOTB ECOS7 STANDARD; PRT; 396 AA.
ID SOTB ECOS7
AC P58529;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sugar efflux transporter.
GN SOTB OR Z2173 OR ECGS2135.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller D.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsuno E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Gasegawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Involved in the efflux of sugars. The physiological role
CC may be the reduction of the intracellular concentration of toxic
CC sugars or sugar metabolites. Transports L-arabinose and to a
CC lesser extent IPTG. Seems to contribute to the control of the
CC arabinose regulation (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- SIMILARITY: Belongs to major facilitator superfamily. SotB
CC (TC 2.A.1.2) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: AE005352; AAG56236.1; -
 DR EMBL: AP002557; BAB35558.1; -
 DR PIR: G90895; G90895.
 DR PIR: H85721; H85721.
 DR HAMAP: MF_005177; 1.
 DR InterPro: IPR007114; MFS.
 DR InterPro: IPR005828; Sub transporter.
 DR Pfam: PF00083; sugar tr; 1.
 DR PROSITE: PS50850; MFS; 1.
 KW Transport; Sugar transporter; Transmembrane; Inner membrane;
 KW Complete proteome.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 50 70 POTENTIAL.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 103 123 POTENTIAL.
 FT TRANSMEM 136 156 POTENTIAL.
 FT TRANSMEM 170 190 POTENTIAL.
 FT TRANSMEM 209 229 POTENTIAL.
 FT TRANSMEM 246 266 POTENTIAL.
 FT TRANSMEM 275 295 POTENTIAL.
 FT TRANSMEM 299 319 POTENTIAL.
 FT TRANSMEM 333 353 POTENTIAL.
 FT TRANSMEM 364 384 POTENTIAL.
 SQ SEQUENCE 396 AA; 42568 MW; CB6A209B0B6D4F0 CRC64;
 Query Match 4.4%; Score 104; DB 1; Length 396;
 Best Local Similarity 18.8%; Pred. No. 1;
 Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;
 QY 31 VTGCVIGSGDFAKLITRLRCGYHVVIGSRNPFSEFPVVDVTHEDALTKNIIF 90
 Db 31 VPGVGL--SDIAQPFHQTAQVGMILTYAWVVALMSLPFLMTSQRERKLLCLEVVF 88
 QY 91 VAIRREHTSLMDLRLHVLGKILLDVSNRRINQYSPESNAEYLASLPDSLYKGVVVS 150
 Db 89 IASHVLSTLS-WSTFTVLISRI-----GVAFAHAI-----WSTTA 123
 QY 151 AMALQGLP--KDAERQVYICSNNIQARQVIELAR-----QLNFIPIDLGSL----- 195
 Db 124 SLAIRMAPAGKRAQALSLIAGTALMAWLGFLGRIVQYQGWMTFFALIGLITLLC 183
 QY 196 -----SSAREIENLPRLFTLMRGVVAISLAT-----FFFLISFVADVHPY 239
 Db 184 LLIKLPILPSEHSGSLKSLPL-----LFRPALMSIYLLTVVVTAAHYTASY-----LEPF 235
 QY 240 ARN---QOSDFPKPIEIVNKTLPVIAITLLSLVYLAGLAAAYQLYGTGRRPFWLE 296
 Db 236 VQNIAGFSANF-----ATALLLLCGAGIIGSVIFGLGNQYAS----- 274
 QY 297 TWLOCRKQLGLSFFPAAVHVAYSICLPFRSERLYFLNMAVQCVHANINSENEEVR 356
 Db 275 -----ALVSTAILLVCLMLLPANSE-----IHGVLSIF-----WG 309
 QY 357 IEMYSFGIMSGILST-----LAVTSPISVSN-----ALNRRPSFIQST 397
 Db 310 IAAWITIGCMQVXVATLAPDADVMALPSGIFNIGICAGALVGNQSLHMS-----MSM 364
 QY 398 LGYVALLISTFHVLYG-----WGRAFEEE 422
 Db 365 IGYVG-TVPAPALIMSIIFRRPMTLEEQ 394
 RESULT 12
 SOTB_ECOLI STANDARD; PRT; 396 AA.
 AC P31122; P76883; P77353;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Sugar efflux transporter.
 GN SOTB OR B1528.

OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCB1_TaxID=562;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474 (1997).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12
 RX MEDLINE=9725157; PubMed=9097039;
 RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sampaio G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horichi T.,
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28,040,1 min region on the linkage map";
 RL DNA Res. 3:363-377 (1996).
 RN (3)
 RP SEQUENCE OF 217-396 FROM N.A.
 RX MEDLINE=93186717; PubMed=838113;
 RA Cohen S.P., Haechter H., Levy S.B.;
 RT "Genetic and functional analysis of the multiple antibiotic
 resistance (mar) locus in Escherichia coli";
 RL J. Bacteriol. 175:1484-1492 (1993).
 RN (4)
 RP CHARACTERIZATION.
 RC STRAIN=SB0;
 RX MEDLINE=99194728; PubMed=10094697;
 RA Bost S., Silva F., Belin D.;
 RT "Transcriptional activation of ydeA, which encodes a member of the
 major facilitator superfamily, interferes with arabinose accumulation
 and induction of the Escherichia coli arabinose PBAD promoter";
 RL J. Bacteriol. 181:2185-2191 (1999).
 RN (5)
 RP CHARACTERIZATION.
 RC STRAIN=US219;
 RX MEDLINE=99369894; PubMed=10438792;
 RA Carole S., Pichoff S., Bouche J.-P.;
 RT "Escherichia coli gene ydeA encodes a major facilitator pump which
 exports L-arabinose and isopropyl-beta-D-thiogalactopyranoside";
 RL J. Bacteriol. 181:5123-5125 (1999).
 CC - FUNCTION: Involved in the efflux of sugars. The physiological role
 may be the reduction of the intracellular concentration of toxic
 sugars or sugar metabolites. Transports L-arabinose and to a
 lesser extent IPTG. Seems to contribute to the control of the
 arabinose regulon.
 CC - SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 (Probable).
 CC - SIMILARITY: Belongs to major facilitator superfamily. SotB
 (TC 2.A.1.2) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000250; AAC74601.1; -
 DR EMBL: D90795; BAA15210.1; -
 DR EMBL: D90796; BAA15218.1; -
 DR EMBL: D90797; BAA15230.1; -


```

Db      306 YLLAYIGIILAGQNSDFELTRK-----ISIVYRKLFTTLGSPCVIPIMCLLYISXN 359
Qy      327 RSEBYELFNAYQCVHANINSENNEEVRLEM---YISF-GIMSL-----GLSLILAY 376
Db      360 FYSTVIFELTLA---NSTLSFSYCGQLINALDIAPRYGFKAVTALIGMGLISSTLA 415
Qy      377 TSISVSNAALNMRSEFSIGTLGYALLISFHTL-----IYGMKR 417
Db      416 GLIINQDFEYAMHKISFLMAGINVTCLV---FYFLPAKGEIQDMAX 458

RESULT 14
NUCC_NEPOL STANDARD: PRT, 391 AA.
AC Q9TRK6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NAD(P)H-quinone oxidoreductase chain H, chloroplast (EC 1.6.5.-)
DE (NAD(P)H dehydrogenase, chain H) (NADH-plastoquinone oxidoreductase
DE 49 kDa subunit).
GN NDH.
OS Nephroselmis olivacea.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;
OC Chlorodendraceae; Chlorodendraceae; Nephroselmis.
OC NCBI_Taxid=31312;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=NIES-484;
RC MEDLINE=99398694; PubMed=10468594;
RA Turnel M., Otis C., Lemieux C.;
RT "The complete chloroplast DNA sequence of the green alga Nephroselmis
RT olivacea" insights into the architecture of ancestral chloroplast
RT genomes."
RT Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253 (1999).
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE PLASTOQUINONE. COMPONENT OF THE IRON-SULFUR (IP) FRAGMENT OF
CC THE ENZYME.
CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) +
CC plastoquinol.
CC -1- SIMILARITY: Belongs to the complex I 49 kDa subunit family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: AF137379; AAD54891.1;
DR InterPro: IPR001135; Oxidored_49kDa.
DR Pfam: PF00346; complex1_49kd_1.
DR PROSITE: PS00535; COMPLEX1_49K; 1.
KM Oxidoreductase; NAD, NADP, Quinone; Plastoquinone; Chloroplast.
SQ SEQUENCE 391 AA; 44913 MW; 9A40AE6895825E CRC64;

Query Match 4.4%; Score 102.5; DB 1; Length 391;
Best Local Similarity 17.9%; Pred. No. 1.3; Indels 173; Gaps 20;
Matches 80; Conservative 70; Mismatches 124;

Qy      85 KNIIVAIHREHYTLMDRLHY---GKILID-----VSNMNRINOYE 127
Db      5 KIDPMIVSM-GPHHPNMGVLRILVLDGNAVLDCEPVGVYLRHGMKKAERKTIQYIP 63
Qy      128 --SNAEYLAFLPDSLIYKGFNVSAWALQLGPKDASROYIICSNNIQARQVIELARQL 185
Db      64 YVTRMVDYLAATFEATIV-----NAPERLANIEEVRRA 96
Qy      186 NFPIPIDGLSSAREIENLPLRLFTLMRGVVAIALATFEFLYSVVRVPHYVANNQS 245

```

```

Db      97 SYLRVIMLELR-----IASHLMLGPFMADLSAQTPPEFYILREEMTY----- 140
Qy      246 DPKYPIELVNTLEPIVALITLSLYLAGLAALQYGVKRYRFPFMLETWLQ-CRKO 304
Db      141 DLFE-----AATGRMAMENFRVGVAA-DVRYG-----WIKKDLFCERY 180
Qy      305 LGILSFFFMV-----HVAYSICLPM-----RSEERYFL 334
Db      181 LPKVDEYELLIRNPILFKRVKVGCTISPOQAINWGLSGPMLRASGVSMDLRKVDRY--- 237
Qy      335 NMAVQCVHANINSENNEEVRARI-----EMYSFGIMSLGLSLILAYSIIP----- 380
Db      238 -EYEDPFHWSVSEETGCLARYLVIREMRTSTYIVQ-----QALSTIRGEPENIEA 290
Qy      381 ---SVSNALNMRSEF-----IOSTLGYVAL-LISTFVLLIYGW 415
Db      291 ROLSGRSTSPNMFYQVFLGKASPTFKMPPRQEHVRYEARPGELGVFLIGDDHVPWRW 350
Qy      416 KQAFEEYRYFTPPNFVIALVLSIV 442
Db      351 K-----IRPGFINVQILPMLV 367

RESULT 15
Y304_BRUME STANDARD: PRT, 220 AA.
ID Y304_BRUME
AC Q9YD73;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein BME110304.
GN BME110304.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OC NCBI_Taxid=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISM / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA Delvecchio V.G., Kaparatzi V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Resnik G.,
RA Jablonski L., Larsen L., D'Souza M., Bernal A., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselcorn R., Kyrides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RT Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the UPF0191 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: AF009669; AAL53546.1;
DR PIR: AG3547; AG3547.
DR HAMAP: MF_01207; -1.
DR InterPro: IPR007916; UPF0191.
DR Pfam: PF05252; UPF0191; 1.
KM Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 20
FT TRANSMEM 39
FT TRANSMEM 54
FT TRANSMEM 72
FT TRANSMEM 85
FT TRANSMEM 104
FT TRANSMEM 124
FT TRANSMEM 146
FT TRANSMEM 153
FT TRANSMEM 175
FT TRANSMEM 179
FT TRANSMEM 198
SQ SEQUENCE 220 AA; 24815 MW; 182C0244743B17FA CRC64;

```

```

Query Match      4.3%; Score 102; DB 1; Length 220;
Best Local Similarity 23.8%; Pred. No. 0.72;
Matches 43; Conservative 25; Mismatches 51; Indels 62; Gaps 7;

QY 267 LSIYVYLAGLLAAAYQLYYGT-----KYRFRPFWLETW-----238
Db 20 LMLIYTYGVPVAVNTFTYGATGQAGADPVKTFEHLGJMALRFLILTLLVTPMRDLTGIT 79
QY 299 -LQCRKQGLISFPFAMVAVAYSLCLPFRSERYLFINNAVQOVHANINENSNNNEEYWR 357
Db 80 LIRYRALGGLAFYVALMHT-----TYVLDQGLNLSAIIIDIVR- 120
QY 358 EMYISFGIMSGLSLAVTSIPSVSNLNV-----REFSFIQSTIGYVALISTFHV 411
Db 121 RPFITIGMISLALLVPLALTS-----NNWSIRKLGRRWSSLHK-LVYTAIGSAVHFL 172
QY 412 I 412
Db 173 M 173

```

Search completed: April 16, 2004, 12:39:41
 Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:37:02 ; Search time 21 Seconds

(without alignments)
2079.569 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351
Sequence: 1 MESTSMGSKSLSETCPN.....ALVLPSTVLDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_78:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160	6.8	198	D95285	conserved hypochet
2	159.5	6.8	211	AC2560	hypochetrical prote
3	156.5	6.7	239	T50571	probable oxidoredu
4	147	6.3	213	H69400	conserved hypochet
5	142	6.0	212	D69351	conserved hypochet
6	142	6.0	224	T10120	P20-dependent NAD
7	133	5.7	223	D64467	hypochetrical prote
8	121	5.1	191	F86826	hypochetrical prote
9	120.5	5.1	232	A69131	conserved hypochet
10	119.5	5.1	216	T00121	hypochetrical prote
11	116.5	5.0	242	G82642	conserved hypochet
12	116	4.9	198	AB3182	conserved hypochet
13	111.5	4.7	695	JR0898	folilitropin recept
14	111.5	4.7	1228	S59681	probable membrane
15	110.5	4.7	222	B84410	hypochetrical prote
16	110.5	4.7	694	JC4301	folilitropin recept
17	109	4.6	320	T28379	ORF MSV218 hypoch
18	109	4.6	712	S50969	probable membrane
19	109	4.6	1184	H71436	hypochetrical prote
20	109	4.6	1301	D65188	disease resistance
21	107.5	4.6	629	S60385	probable membrane
22	107.5	4.6	695	Q8H0FT	folilitropin recept
23	105	4.5	442	B64582	sodium- and chlori
24	105	4.5	604	T11042	hypochetrical prote
25	104	4.4	396	C64907	chloramphenicol re
26	104	4.4	396	G90895	hypochetrical prote
27	104	4.4	396	H85721	probable resistance
28	103	4.4	465	S69315	sodium-phosphate t
29	102.5	4.4	314	AB2972	hypochetrical prote

30	102.5	4.4	314	2	H99310	probable peptide A
31	102.5	4.4	1242	2	T39453	probable mra stab
32	102	4.3	220	2	AG3547	bicyclomycin resis
33	102	4.3	574	2	T41068	hypochetrical prote
34	101	4.3	320	2	E71139	hypochetrical prote
35	101	4.3	348	2	T12284	NADH2 dehydrogenas
36	101	4.3	420	2	F63144	O-antigen transpor
37	101	4.3	501	2	T02134	hypochetrical prote
38	101	4.3	503	2	C86250	hypochetrical prote
39	101	4.3	735	2	AB3006	hypochetrical prote
40	100.5	4.3	346	2	T11181	NADH2 dehydrogenas
41	100.5	4.3	1780	2	AB5045	probable glucan sy
42	100	4.3	395	2	C71219	hypochetrical prote
43	100	4.3	476	2	A28439	endonuclease ScelI
44	100	4.3	714	2	AF2479	ABC transporter Ar
45	99.5	4.2	452	2	C71391	NADH2 dehydrogenas

ALIGNMENTS

RESULT 1
D95285
Conserved hypothetical protein Sma0349 [Imported] - Sinorhizobium meliloti (strain 1021)

C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C/Accession: D95285

R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows, R.; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yen, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A/Reference number: A95262; MUID:21396509; PMID:11481432

A/Accession: D95285

A/Molecule type: DNA

A/Residues: 1-198 <KUR>

A/Cross-References: GB:AB006469, PIRN:AAK64846.1, PID:G14523260, GSPDB:GN00165

A/Experimental source: strain 1021, megaplasmid pSymA

R/Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, A.;

heault, P.; Vandenbol, M.; Vorholter, F.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.

A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A/Reference number: A96039; MUID:21368234; PMID:11474104

A/Contents: annotation

C/Genetics:

A/Genome: plasmid

Query Match 6.8%; Score 160; DB 2; Length 198;
Best Local Similarity 26.7%; Pred. No. 2.1e-05;

Matches 55; Conservative 40; Mismatches 83; Indels 28; Gaps 8;

QY	32	TVGVIGSGDFASLTIRLRGCVHVI--GSRNPKAS---EEFRVIVVTTHEDALTKTN	87
DB	3	TTAIIAGAGIGALMERFTAAQIPAIANSKRPASLSVTFDFGASVAVEIKDL--QAD	61
QY	88	IFVVAIHREHYTSLMDLRHLV---GKILIDVSNMRINQYP-----ESNAEYASLFP	138
DB	62	VVLAV---PYDSIADIVTVQVSDWCGQIVDPASNAIDPPAFKPRDLGRLSTEIVSEIVP	118
QY	139	DLIVKGFNVASAMQLQPK--DASROYICSNNTQARQOVIETARLNFPIDLGSLT	196
DB	119	GAKVVAKEFTLPAAVLADPEKGTGSRVLFSGNSDANROVAELISLGRFVLDGLTA	178
QY	197	SAREIENPLRLFTLRGFWVAISL	222
DB	179	ASGPICQF-----GRPLVALNT	195

RESULT 2

AC2560
 hypothetical protein alr8074 [imported] - Noctoc sp. (strain PCC 7120) plasmid pCC7120g4
 C:Species: Noctoc sp. PCC 7120
 A:Note: Noctoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AC2560
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurlitz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
 A:Reference number: AB1807, MUID:21595285, PMID:11759840
 A:Accession: AC2560
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-211 <KUR>
 A:Cross-references: GB:AP03603; PIDN:BA87404.1; FID:gl7134847; GSPDB:GN00182
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr8074
 A:Genome: plasmid

Query Match 6.8%; Score 159.5; DB 2; Length 211;
 Best Local Similarity 24.9%; Pred. No. 2.4e-05;
 Matches 51; Conservative 43; Mismatches 70; Indels 41; Gaps 8;

QY 33 VGVIGSGDFAKSLTIRLCGYHVVIGSRP-----KFASEFPFHVVDVTHHEDALTKTN 87
 Db 3 IAFIGIGVGSGALASQSLSDHTVTIARNSNSDVTKALAKYPE-LQVSSPDAIAQAE 61
 QY 88 IIFVA---IHREHYTSLMDRLHVLGKILIDVGNRRIN-----QYBESNAEVLASLP 138
 Db 62 VIFLATPTTANQALAEVGLDS---GKILVDCINPGVALTHLKEQSGSELVGSFVP 117
 QY 139 DSLIVKGFNVASVALQIGPDASROY-----ICSNNIQARQVIELARQINF 187
 Db 118 HAKVYKAFIT-----YGFENFENHTYPGYGNLKPALVAGNDVPKQVSTLCQQLGW 170
 QY 168 IPIDGSSASAREIENLPLRLFTLM 212
 Db 171 EAVDVGNLSMGLHEHMTL-----LM 191

RESULT 3
 T50571
 probable oxidoreductase [imported] - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 28-Jul-2000
 C:Accession: T50571
 R:Redenbach, M.; Kleiser, H.M.; Denapalte, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw, M.
 Mol. Microbiol. 21, 77-96, 1996
 A:Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S
 A:Reference number: Z20556; MUID:97000351, PMID:8843436
 A:Accession: T50571
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-239 <RED>
 A:Cross-references: EMBL:AL33220; PIDN:CAB61708.1
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Note: SCC75A.08c
 C:Superfamily: conserved hypothetical protein M1501

Query Match 6.7%; Score 156.5; DB 2; Length 239;
 Best Local Similarity 27.9%; Pred. No. 5e-05;
 Matches 61; Conservative 41; Mismatches 72; Indels 45; Gaps 10;

QY 25 IKDKARKVVGVI--SGGDFAKSLTIRLCGYHVVIGSR---NPKFASEFPFHVVDVTHHE 80
 Db 23 LPDVSGLVVGVGIGPQGGKGLAYRLAKAGKATVIGSFAARAAAABEIGHGVGADNA 82
 QY 81 DALTKNIIFAIHRE-HYTSIMDLRLHVLGKILIDVGNRRINQ-----YBP--SNAE 131
 Db 83 ETARSDVIVAVPMDGKLTSLRLAELSGKLVDVCPNPGFGKGAVALKPEEGSAAE 142

QY 132 YLASFPDSLIVKGFNVASVALQIGPDASROYICSNNIQARQVIELARQINF-PI 178
 Db 143 QAAALLPDSRAAAHHHSATVLLQDPEIDEDTDMVNGERADVEI-----VGA----- 192
 QY 179 IELARQNLFP---IDGSLSSAREIENLPLRLFTLM 213
 Db 193 --LAQR--IPGMRGVFAGRRLNHAQVESLVANLISVNR 226

RESULT 4
 H69400
 conserved hypothetical protein AF1209 - *Archaeoglobus fulgidus*
 C:Species: *Archaeoglobus fulgidus*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
 C:Accession: H69400
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: H69400
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-213 <KLE>
 A:Cross-references: GB:AE001021; GB:AE000782; NID:92689344; PIDN:AA890038.1; PID:9264937
 C:Superfamily: conserved hypothetical protein M1501

Query Match 6.3%; Score 147; DB 2; Length 213;
 Best Local Similarity 27.5%; Pred. No. 0.0024;
 Matches 56; Conservative 44; Mismatches 68; Indels 36; Gaps 11;

QY 37 GSGDFAKSLTIRLCGYHVVIGSRN-----KFASEFPFHVVDVTHHEDALTKNI 88
 Db 8 GTGNIGEGALMWGLGVEIIVGSRKLEKAEIADYIKVGDASITGMREDAEATCD- 66
 QY 89 IYVAIHREHYTSLMD-----LHLLVGLKILID-----VSNMNRINQYF--SNAE 131
 Db 67 --VAV---FTIPWEFADTAEMLKROLAGKVIISPLVPMKKVGNF--VYVRPEEGSAE 119

QY 132 YLASFPDSLIVKGFNVASVALQIGPDASROYICSNNIQARQVIELARQINF-PI 190
 Db 120 KLASVTEBSVVAHAHSIPARFALGEEFENDVTCDS--CAKEVYVDLTKISGLAAL 178

QY 191 DLGSLSSAREIENL-PLRLFTLM 213
 Db 179 DAGLSNHLVSTLPLILNWK 202

RESULT 5
 D69361
 conserved hypothetical protein AF0892 - *Archaeoglobus fulgidus*
 C:Species: *Archaeoglobus fulgidus*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
 C:Accession: D69361
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: D69361
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-212 <KLE>
 A:Cross-references: GB:AE001042; GB:AE000782; NID:92689365; PIDN:AA890348.1; PID:9264970
 C:Superfamily: conserved hypothetical protein M1501

A:Reference number: A63400; MUID:96337999; PMID:8688087
 A:Accession: D6447
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-223 <full>
 A:Cross-references: GB:U67591; GB:L77117; NID:92826422; PIDD:AA99514.1; PID:G1500389; T
 C:Genetic:
 A:Map position: REV1473617-1472946
 C:Superfamily: conserved hypothetical protein MJ1501

```

Query Match      5.7% ; Score 133 ; DB 2 ; length 223 ;
Best Local Similarity 21.9% ; Pred. No. 0.0033 ;
Matches 46 ; Conservative 52 ; Mismatches 78 ; Indels 34 ; Gaps 7 ;

QY 37 GSGDPAKSLTIRLIRCGYHVVIGSRNPKAFSEPPPHVVDV-----THEDALT 84
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8 GTGQDGGGLRLRLAK--NNKIIIGSRKKEKEEAKAKKELLKRGIEADIIIGENDDAAK 66
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 85 KTNIIIFAIRRH--YTSLMDLRHLVGLKILIDV-----SNMRINQYPE-SNAEYL 133
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 67 EGDVVIISLEYEYTLSTIKDLKEELKQIVASIGVPLATIGKPRRLLEPPPGSAVEMV 126
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 134 ASLEPDSLIVKGFNVVSAMALQGPQDASQVYICSNNIQARQQVTELAQNLFI-PIDL 192
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 127 QNVLKESKRVVSAPQNVCHAVLEDDLDNPEVDDIIVCGNDKAEKRVIDLAMQIDGVRAIDC 186
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 193 GSLSARETENLPLRLFTLWRGPVVAISL 222
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 187 GNLEKSKIIIEAIT-----PULIGINT 207
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 8
F86826
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #next_change 03-Aug-2001
C:Accession: F86826
R:Botolin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A:Reference number: A86625; MUID:2123516; PMID:11337471
A:Accession: F86826
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-191 <SFO>
A:Cross-references: GB:E005176; PTD:G12724621; PINN:AAK05712.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ygfE

```

```

Query Match      5.1%; Score 121; DB 2; Length 191;
Best Local Similarity 22.2%; Pred. No. 0.024;
Matches 47; Conservative 33; Mismatches 74; Indels 58; Gaps 7;

QY      32 TVGVIIGSGDFPKSLTITLIRGCHVYIGSKNPKFASFFPHVYDYTHEDALTKTN---- 87
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db       3 TISIFGRKGKQKGA-----GDNFSSSVKNKNTILSSKTELGEI 42
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      88 -----IIFVAIHREHYTSLMDRLHLVGKILIDVSNNRINQYPR-----SNAEYL 133
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      43 VVLAVPYVALAGITQGEVSTD-----LQGIILIDINPVDFTFDSLLVPSDTSAALAI 95
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     134 ASLSPDSLLTYGFRVNVASAMALQLGPKDASQ-----VYICNNNIQARQVTELMARQLNFI 189
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      96 AKQLPNSMTIVAEN--TFPSDTLATKRVAAHQGTVVLASDSQAEKETITIALENGSL 153
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     190 IDLGSLSSAREIENTPLRLPTL-----WRG 214
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     154 LDAGSLKRARELEAIGFQITFLASEKISWDG 185
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

conserved hypothetical protein MTH248 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C:Accession: A69111
R:Smith, D.R.; Docetette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qu, D.; Spatafora, R.; Vicaire, C.J.; Mao, J.; Rice, P.; Noelling, J.; Seever, J.N.
X1, S.; Church, G.M.; Daniels, C.U.;
U: Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:93711463
A:Accession: A69111
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-232 <MTH>
A:Cross-references: GB:AE000811; GB:AE000666; NID:g6621287; PIDN:AA84754.1; PID:g9262129
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH248
A:Start codon: TTG
C:Superfamily: conserved hypothetical protein M01501

RESULT 10
T00121
hypothetical protein 8 - Leptospira interrogans
C|Species: Leptospira interrogans
C|Date: 22-Jan-1999 #sequence-revision 22-Jan-1999 #text_change 21-Jul-2000
C|Accession: T00121
R|Takahashi, Y.; Akase, K.; Hirano, H.; Fukunaga, M.
Gene 215, 37-45, 1998
A|Title: Physical and genetic maps of the Leptospira interrogans serovar Icterohaemorrhagiae
A|Reference number: Z14115; MUID:98332717; PMID:9666070
A|Accession: T00121
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-216 <Tak>
A|Cross-references: EMBL:AB010203; NID:g2780763; PIDN:BA424368.1; PID:g2780765
A|Experimental source: strain Ictero No.1, substrain icterohaemorrhagiae

Dd	126	APAAAVKFFSVGS	GLMVPQLKGEKPSNF	ICGNDSSKKQIKELIDT	FGWTEDMGK	184
Qy	195	LSGARREINLPRLFTLW	212			
Dd	185	VEAARAIEPLCT---	LM	198		

```

RESULT 11
GS2642
Conserved hypothetical protein XFI737 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #ext_change 20-Aug-2000
C:Accession: GS2642
R:anonymous, the Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: GS2642
A:Status: preliminary

```

```

RESULT 12
A3182
conserved hypothetical protein Atu183 [imported] - Agrobacterium tumefaciens (strain C
C.Species: Agrobacterium tumefaciens
C.Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 16-Nov-2002
C.Accession: A3182
R.Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
; Karp, P.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McCellen
Science 294, 2317-2323, 2001

```

Query Match 4.7%; Score 111.5; DB 1; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.71; Mismatches 151; Indels 183; Gaps 16;
Matches 84; Conservative 75;

98 HEIRIEKANNL-LYINPEAFQNLPNLRILLISNTGKHLPDVYKHISFGKVLIDQDNIN 156
122 INQVPESSNAEYLASLPFDSLTIVKGFVNVSAWALQCPKASRQVYICSNNTQARQVIEL 181
157 IH-----TIERNSEFVGISFESVILWL-----NKNQGIQIEHNCA----- 189
182 AROLNFIPIDGSISSARIEINLPRLFTLMRGVVAISLATFFFLYGFVHDVHPVAR 241
190 ---FNGQIDELNDSNNNEELPNDVFGHASPVLIDISRRIRHSLPSYGLENLKRA 246
242 NQSDPFKITEIVNKLPLVIAITLISLVY----- 271
247 RSTYNLKKLP-----SLKKVLAEMEASLTYPHSCAFANMRQISLHDI CNKSLIRQEV 301
272 -----TAGLLAAVQLYGYGKRRF-----PMTETMLQCKQKG 306
302 DYMQTGGQSSLAEDNESSYSGFDMTYAFEDYDLCNEVVDTCSPKRDANPCDIDIG 361
307 ----LISFFPAMVHA-----YSICLPMRBSERYLFNNAYQQVYHANIENS 348
362 YNIRLVIMFISILAIIGNIIVLVTTSQYKLTVP-----RLMCNLAIFAD----- 408
349 WNEEVRRIEMVYSIFGIMSLGLSLAVTSIPVSNAIMRE-----FSFIQSTLG 399
409 -----LCIGYLLILIASVDIHNTSQYHNAAIDMQTAGCDAGFFVFGSELS 456
400 --VVALISTFH-----VLIVGKKRAFEERYRFTYPNV-LAL 436
457 VYTLATLTERRHITTHAMQDCKVHVAASVWVGWIFAFAPALFPIFGISSYMKVSI 516
437 VLPSVILDLQL 449
517 CLPMDIDSPLSQL 529

Db

RESULT 14
559681
probable membrane protein YPL012w - Yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein LPAW; hypothetical protein YPL012.01
C/Species: Saccharomyces cerevisiae
C/Date: 13-Jan-1996 #sequence-revision 01-Mar-1996 #text-change 19-Apr-2002
C/Accession: S59681; S52519
R/Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; Wa
submitted to the EMBL Data Library, August 1995
A/Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.
A/Reference number: S59677
A/Accession: S59681
A/Molecule type: DNA
A/Residues: 1-1228 <HALL>
A/Cross-references: EMBL:U03335; NID:G965076; PID:G965081; MIPS:YPL012w
R/Baddock, K.; Churcher, C.
submitted to the EMBL Data Library, February 1995
A/Reference number: S52519
A/Accession: S52519
A/Molecule type: DNA
A/Residues: 220-1228 <BAD>
A/Cross-references: EMBL:Z48483; NID:G683777; PID:G683778; MIPS:YPL012w
A/Experimental source: strain AB972
C/Genetics:
A/Gene: SGD:RP12
A/Cross-references: SGD:S0005933
A/Map position: 16L
C/Keywords: transmembrane protein
F/125-141/Domain: transmembrane #status predicted <TM1>
F/478-494/Domain: transmembrane #status predicted <TM2>
F/1720-736/Domain: transmembrane #status predicted <TM3>

Query Match	4.7%;	Score 111.5;	DB 2;	Length 1228;
Best Local Similarity	19.9%;	Pred. No. 1.5;		
Matches 111;	Conservative 71;	Mismatches 152;	Indels 225;	Gaps 28;

```

Dy      11 KSIASETLPGINIKDARK-----VTVGVISSGDFASLITRLI 50
      292 KTMAEITISSGLAENKTLRVLDITFALKPNSVDTLTKRSIAVIVKMSYVAHHPKXAL 351
Qy      51 R-----CGYHVIVGSRNKF--ASEFFPVVDVTHEDAL----- 83
Dy      352 RKJPGVEHIMCTY---LASETFEYVOAASQCLISILSESVKDDLTYTPSVDKRVKND 408
Qy      84 -----TKTMIIFAVIREHYTSLMDRLHLVGKILIDVSNMKNINQYPSNAEYLASLF 137
Dy      409 EIISQAKTFIDFISTRYSHCS-----REIL--KILVAANKPRY---RSNPHFLKSL- 456
Qy      138 PDSILIVGFNVVSAWALQLGPKDASROYICSNNOARQOYIELAROLNTI----- 188
Dy      457 -----KIVDWVRV-----EEOFMDRNEIEIVIGASISAMG 488
Qy      189 -----PIDIGLSGARBIENPLRLFTLMRGPV--VAISLTFEFLYSFYVDVI 236
Dy      489 PEMILAEAPLNDPSSER-----PGR--AMLLPLRDYKXANNALTF-----QNEL 533
Qy      237 HPVARNQOOSDYKPIPIELVN-----KLP-----IYAITLSTL 269
Dy      534 APFYKSSQSKDKPESIEIQLRVQITVDQIMWSTLPRCELPMDLRASFDFEFSSELSTL 593
Qy      270 VYL-----AGLAAAYQLVYGTK-----RRFP-----PMLETWLOCRK 303
Dy      594 LYSEVELRTTICHALKVLAESNVSAEBSHHVLLLRPFISAEQKNIETLS-----K 648
Qy      304 QLGILSFFPMVAVAYSICLPMRS-----ERL-----FLMAYQOYHANIENSNWE 351
Dy      649 STNLLAALFNV-----YTQITPMARSYLLETIDQLYKTSKEDKERTNNVCGLKSNMNE 704
Qy      352 EEWVRIEYVIFGIMSLGLSTL--AVTSIPSVSNALNMBRESEFIOSTGLGVALLSTPH 409
Dy      705 ESSGNVNMKEKKPQLTATLDDLIICMITLVPSYSALFMSF-----LTVNSAD 754
Qy      410 VLIYGMKRAFEERYFRYT 428
Dy      755 ALIQ--KRA-----YRIIT 766

```

RESULT 15
B84410
Hypothetical protein Vng2607c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
A:Accession: B84410
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Lechauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Madocks, D.G.; Cabid
Jung, K.H.; Alam, M.; Fedkes, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: AB4160; MUID:20504483; PMID:11016950
A:Accession: B84410
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-222 <STO>
A:Cross-references: GS:AE004437; NID:GI0581987; PIDN:AMG20646.1; GSFDB:GN00138
C:Genetics:
A:Gene: VNG2607C
C:Superfamily: conserved hypothetical protein M01501

Query Match	4.7%;	Score 110.5;	DB 2;	Length 222;
Best Local Similarity	24.3%;	Pred. No. 0.2;		
Matches 49;	Conservative 34;	Mismatches 94;	Indels 25;	Gaps 6;

```

Db      8 GTGDDGAGLALWATDSDHDI.VIGSRDEKARFETAAAYEDTLADQGVDRKLTGPAENMAA 67
Ox      84 TKTNITFVAIHEHETSLM-----DLRHLLVGKILDVSNMKRIQVPSNAE---YL 133
Db      68 DRADVAVVAVPAHYHTDVWAGVADLDDLTIVSPANGIAGGEGHTNPPSAGSVTALV 127
Ox      134 ASLFPDSL-IVKGFNVVSAHALQLGPKDASHQVYICSNNIQARQVITELARQLNFT-PID 191
Db      128 ADAAPRGVGVVGAFFHNLADRLADDTETLDDLTIVGNDGEGARTRVAELADDTIGLRALD 187
Ox      192 LGSLSAREIENLPRLFTLMR 213
Db      188 AGPVENAAEVESITPLILTLAR 209

Search completed: April 16, 2004, 12:41:12
Job time : 23 secs

```

QY 37 GSGDFAKSLTRLRCGYH-VIGSRNPKFASFPVVDVT-----HHEDAL 83

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:40:42 ; Search time 48 Seconds
(without alignments)
2607.409 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351
Sequence: 1 MBSISMMGSKSLSTFCPLP.....ALVPSIVLLDLQLCRPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2351	100.0	454	10	US-09-888-257A-10 Sequence 10, Appl
2	2351	100.0	454	12	US-10-455-822-3 Sequence 3, Appl
3	2351	100.0	454	12	US-10-455-822-7 Sequence 7, Appl
4	2351	100.0	454	12	US-10-455-822-9 Sequence 9, Appl
5	2351	100.0	454	12	US-10-455-822-19 Sequence 19, Appl
6	2351	100.0	454	12	US-10-455-822-21 Sequence 21, Appl
7	2351	100.0	454	12	US-10-455-822-23 Sequence 23, Appl
8	2351	100.0	454	12	US-10-455-822-25 Sequence 25, Appl
9	2351	100.0	454	12	US-10-455-822-31 Sequence 31, Appl
10	2351	100.0	454	12	US-10-455-822-33 Sequence 33, Appl
11	2351	100.0	454	12	US-10-455-822-35 Sequence 35, Appl
12	2351	100.0	454	12	US-10-455-822-37 Sequence 37, Appl
13	2351	100.0	454	12	US-10-455-822-39 Sequence 39, Appl
14	2351	100.0	454	12	US-10-455-822-78 Sequence 78, Appl
15	2351	100.0	454	12	US-10-455-822-121 Sequence 121, Appl

16	2351	100.0	454	12	US-10-455-822-160 Sequence 160, App
17	2351	100.0	454	12	US-10-455-822-161 Sequence 161, App
18	2351	100.0	454	12	US-10-455-822-162 Sequence 162, App
19	2351	100.0	454	12	US-10-455-822-166 Sequence 166, App
20	2351	100.0	454	12	US-10-455-822-167 Sequence 167, App
21	2351	100.0	454	12	US-10-455-822-168 Sequence 168, App
22	2351	100.0	454	12	US-10-455-822-173 Sequence 173, App
23	2351	100.0	454	12	US-10-455-822-179 Sequence 179, App
24	2351	100.0	454	12	US-10-455-822-185 Sequence 185, App
25	2351	100.0	454	12	US-10-455-822-191 Sequence 191, App
26	2351	100.0	454	12	US-10-455-822-193 Sequence 193, App
27	2351	100.0	454	14	US-10-165-044-8 Sequence 8, Appl
28	2351	100.0	454	15	US-10-239-607-29 Sequence 29, Appl
29	2351	100.0	454	15	US-10-239-607-29 Sequence 29, Appl
30	2351	100.0	454	15	US-10-239-607-29 Sequence 29, Appl
31	2351	100.0	454	15	US-10-192-190-2 Sequence 2, Appl
32	2348	99.9	454	15	US-10-104-047-3155 Sequence 3155, Ap
33	2347	99.8	454	12	US-10-455-822-29 Sequence 29, Appl
34	2340	99.5	454	12	US-10-455-822-27 Sequence 27, Appl
35	2336	99.4	454	12	US-10-455-822-84 Sequence 84, Appl
36	2294	97.6	444	12	US-10-455-822-85 Sequence 85, Appl
37	2294	97.6	444	12	US-10-455-822-88 Sequence 88, Appl
38	2294	97.6	444	12	US-10-455-822-90 Sequence 90, Appl
39	2294	97.6	490	9	US-09-802-520-1 Sequence 1, Appl
40	2294	97.6	490	12	US-10-455-822-13 Sequence 13, Appl
41	2294	97.6	490	12	US-10-455-822-17 Sequence 17, Appl
42	2294	97.6	490	12	US-10-455-822-51 Sequence 51, Appl
43	2294	97.6	490	12	US-10-455-822-53 Sequence 53, Appl
44	2294	97.6	490	12	US-10-455-822-55 Sequence 55, Appl
45	2294	97.6	490	12	US-10-455-822-57 Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-09-888-257A-10
Sequence 10, Application US/0988257A
GENERAL INFORMATION:
Application No. US20030060612A1
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Polakis, Paul
APPLICANT: Smith, Victoria
APPLICANT: Wood, William I.
APPLICANT: Wu, Thomas D.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
TITILE OF INVENTION: TREATMENT OF TUMOR
FILE REFERENCE: P5002K1
CURRENT APPLICATION NUMBER: US/09/888,257A
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: ~~US-09-069,370~~ NO
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: ~~US-09-069,370~~ NO
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 60/499,792 NO
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: US ~~60/499,792~~ NO
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/12011
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US00/04342
PRIOR FILING DATE: 2000-02-18
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439

Pro 23203
disclosed 9/24/2000
60/235451

disclosed

PRIOR FILING DATE: 2000-03-30
 PRIOR APPLICATION NUMBER: PCT/US00/23328
 PRIOR FILING DATE: 2000-08-24
 PRIOR APPLICATION NUMBER: PCT/US00/32678
 PRIOR FILING DATE: 2000-12-01
 PRIOR APPLICATION NUMBER: PCT/US01/06520
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: PCT/US01/06666
 PRIOR FILING DATE: 2001-03-01
 NUMBER OF SEQ ID NOS: 10
 SEQ ID NO 10
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-888-257A-10

Query Match 100.0%; Score 2351; DB 10; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPEKSLSTCLPENGINGIKDARKYTVGVIGSGDPAKSLTIRLRGCVHVIGS 60
 DB 1 MESISMGSPEKSLSTCLPENGINGIKDARKYTVGVIGSGDPAKSLTIRLRGCVHVIGS 60
 QY 61 RNPKEASEFFPHVVDVTHHEDALTKTNIIIVAIHREHYTSLMDLRHLVKGILLIDVSNM 120
 DB 61 RNPKEASEFFPHVVDVTHHEDALTKTNIIIVAIHREHYTSLMDLRHLVKGILLIDVSNM 120
 QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPKASROYICSNNOARQOYIE 180
 DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPKASROYICSNNOARQOYIE 180
 QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPKASROYICSNNOARQOYIE 180
 DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPKASROYICSNNOARQOYIE 180
 QY 181 LARQNFIPIDIGSLSSAREIENPLRLFTLMRGVVAISLAFEPFLSYFVDYIHPIYA 240
 DB 181 LARQNFIPIDIGSLSSAREIENPLRLFTLMRGVVAISLAFEPFLSYFVDYIHPIYA 240
 QY 181 LARQNFIPIDIGSLSSAREIENPLRLFTLMRGVVAISLAFEPFLSYFVDYIHPIYA 240
 DB 181 LARQNFIPIDIGSLSSAREIENPLRLFTLMRGVVAISLAFEPFLSYFVDYIHPIYA 240
 QY 241 RNOQSDFYKIPLEIYNKTLPIVAITLLSLVYAGLAAAYQLYGKTRRPPMLETWLO 300
 DB 241 RNOQSDFYKIPLEIYNKTLPIVAITLLSLVYAGLAAAYQLYGKTRRPPMLETWLO 300
 QY 301 CRKQGLISFFPAMVAVASLCLPWRSEERYLFNMAVQOVANINENSMNEEVRIRIEM 360
 DB 301 CRKQGLISFFPAMVAVASLCLPWRSEERYLFNMAVQOVANINENSMNEEVRIRIEM 360
 QY 301 CRKQGLISFFPAMVAVASLCLPWRSEERYLFNMAVQOVANINENSMNEEVRIRIEM 360
 DB 301 CRKQGLISFFPAMVAVASLCLPWRSEERYLFNMAVQOVANINENSMNEEVRIRIEM 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTLGYVALLISTHVLIIYGMKAF 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTLGYVALLISTHVLIIYGMKAF 420
 QY 421 EBYRFTYTPNFVLAIVLPISVILLDLQLCRYPD 454
 DB 421 EBYRFTYTPNFVLAIVLPISVILLDLQLCRYPD 454

RESULT 2

US-10-455-822-3
 Sequence 3, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raltano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Eld, Pia M.
 APPLICANT: Fairs, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455, 822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-3

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPEKSLSTCLPENGINGIKDARKYTVGVIGSGDPAKSLTIRLRGCVHVIGS 60
 DB 1 MESISMGSPEKSLSTCLPENGINGIKDARKYTVGVIGSGDPAKSLTIRLRGCVHVIGS 60
 QY 61 RNPKEASEFFPHVVDVTHHEDALTKTNIIIVAIHREHYTSLMDLRHLVKGILLIDVSNM 120
 DB 61 RNPKEASEFFPHVVDVTHHEDALTKTNIIIVAIHREHYTSLMDLRHLVKGILLIDVSNM 120
 QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPKASROYICSNNOARQOYIE 180
 DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPKASROYICSNNOARQOYIE 180
 QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPKASROYICSNNOARQOYIE 180
 DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPKASROYICSNNOARQOYIE 180
 QY 181 LARQNFIPIDIGSLSSAREIENPLRLFTLMRGVVAISLAFEPFLSYFVDYIHPIYA 240
 DB 181 LARQNFIPIDIGSLSSAREIENPLRLFTLMRGVVAISLAFEPFLSYFVDYIHPIYA 240
 QY 181 LARQNFIPIDIGSLSSAREIENPLRLFTLMRGVVAISLAFEPFLSYFVDYIHPIYA 240
 DB 181 LARQNFIPIDIGSLSSAREIENPLRLFTLMRGVVAISLAFEPFLSYFVDYIHPIYA 240
 QY 241 RNOQSDFYKIPLEIYNKTLPIVAITLLSLVYAGLAAAYQLYGKTRRPPMLETWLO 300
 DB 241 RNOQSDFYKIPLEIYNKTLPIVAITLLSLVYAGLAAAYQLYGKTRRPPMLETWLO 300
 QY 301 CRKQGLISFFPAMVAVASLCLPWRSEERYLFNMAVQOVANINENSMNEEVRIRIEM 360
 DB 301 CRKQGLISFFPAMVAVASLCLPWRSEERYLFNMAVQOVANINENSMNEEVRIRIEM 360
 QY 301 CRKQGLISFFPAMVAVASLCLPWRSEERYLFNMAVQOVANINENSMNEEVRIRIEM 360
 DB 301 CRKQGLISFFPAMVAVASLCLPWRSEERYLFNMAVQOVANINENSMNEEVRIRIEM 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTLGYVALLISTHVLIIYGMKAF 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMRSEFIOSTLGYVALLISTHVLIIYGMKAF 420
 QY 421 EBYRFTYTPNFVLAIVLPISVILLDLQLCRYPD 454
 DB 421 EBYRFTYTPNFVLAIVLPISVILLDLQLCRYPD 454

RESULT 3

US-10-455-822-7
 Sequence 7, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raltano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Eld, Pia M.
 APPLICANT: Fairs, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455, 822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 7
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-7

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6,1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSDPAKSLTIRLRICGHVYIGS 60
 DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSDPAKSLTIRLRICGHVYIGS 60
 QY 61 RNPKEASEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQPEESNAEYLAIFPDSLIVKGFNVSAVALQGPDAKROYIICSNNTQAOQVITE 180
 DB 121 RINQPEESNAEYLAIFPDSLIVKGFNVSAVALQGPDAKROYIICSNNTQAOQVITE 180
 QY 121 RINQPEESNAEYLAIFPDSLIVKGFNVSAVALQGPDAKROYIICSNNTQAOQVITE 180
 DB 121 RINQPEESNAEYLAIFPDSLIVKGFNVSAVALQGPDAKROYIICSNNTQAOQVITE 180
 QY 181 LARQNLFPIDIGLSASAREIENLPLRLFTLRGPPVVAISLATFFELYSFVDVIHPYA 240
 DB 181 LARQNLFPIDIGLSASAREIENLPLRLFTLRGPPVVAISLATFFELYSFVDVIHPYA 240
 QY 241 RNOQSDFYKPIPIVAKTLPYVATILSLVYLAGLAAAYQYIGTKYRFPFMTETWQ 300
 DB 241 RNOQSDFYKPIPIVAKTLPYVATILSLVYLAGLAAAYQYIGTKYRFPFMTETWQ 300
 QY 301 CRKQGLISFPFAMVAVASLCLPMRSEERYLFNMAVQOVANINENSWNEEWRLEMY 360
 DB 301 CRKQGLISFPFAMVAVASLCLPMRSEERYLFNMAVQOVANINENSWNEEWRLEMY 360
 QY 361 ISFGIMSLGLSLAATSVISVSANLNMRSEFSTIGTGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAATSVISVSANLNMRSEFSTIGTGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EBYRFPYTPPNFVLAIVLPISIVILDLLQCRYPD 454
 DB 421 EBYRFPYTPPNFVLAIVLPISIVILDLLQCRYPD 454

RESULT 4
 US-10-455-822-9

Sequence 9, Application US/10455822
 Publication No. US20040048798A1

GENERAL INFORMATION:

APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.

APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya

APPLICANT: Chalita-Eld, Pia M.
 APPLICANT: Paris, Mary

TITLE OF INVENTION: Nucleic acids and Corresponding Proteins

TITLE OF INVENTION: Encited 98P4B6 Useful in Treatment and Detection of Cancer

FILE REFERENCE: 51158-20016.24

CURRENT FILING DATE: 2003-06-04

PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned

PRIOR FILING DATE: 2002-12-20

PRIOR APPLICATION NUMBER: US60/370,387

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 9
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-9

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6,1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSDPAKSLTIRLRICGHVYIGS 60
 DB 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIIGSDPAKSLTIRLRICGHVYIGS 60
 QY 61 RNPKEASEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKEASEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQPEESNAEYLAIFPDSLIVKGFNVSAVALQGPDAKROYIICSNNTQAOQVITE 180
 DB 121 RINQPEESNAEYLAIFPDSLIVKGFNVSAVALQGPDAKROYIICSNNTQAOQVITE 180
 QY 121 RINQPEESNAEYLAIFPDSLIVKGFNVSAVALQGPDAKROYIICSNNTQAOQVITE 180
 DB 121 RINQPEESNAEYLAIFPDSLIVKGFNVSAVALQGPDAKROYIICSNNTQAOQVITE 180
 QY 181 LARQNLFPIDIGLSASAREIENLPLRLFTLRGPPVVAISLATFFELYSFVDVIHPYA 240
 DB 181 LARQNLFPIDIGLSASAREIENLPLRLFTLRGPPVVAISLATFFELYSFVDVIHPYA 240
 QY 241 RNOQSDFYKPIPIVAKTLPYVATILSLVYLAGLAAAYQYIGTKYRFPFMTETWQ 300
 DB 241 RNOQSDFYKPIPIVAKTLPYVATILSLVYLAGLAAAYQYIGTKYRFPFMTETWQ 300
 QY 301 CRKQGLISFPFAMVAVASLCLPMRSEERYLFNMAVQOVANINENSWNEEWRLEMY 360
 DB 301 CRKQGLISFPFAMVAVASLCLPMRSEERYLFNMAVQOVANINENSWNEEWRLEMY 360
 QY 361 ISFGIMSLGLSLAATSVISVSANLNMRSEFSTIGTGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAATSVISVSANLNMRSEFSTIGTGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EBYRFPYTPPNFVLAIVLPISIVILDLLQCRYPD 454
 DB 421 EBYRFPYTPPNFVLAIVLPISIVILDLLQCRYPD 454

RESULT 5
 US-10-455-822-19

Sequence 19, Application US/10455822
 Publication No. US20040048798A1

GENERAL INFORMATION:

APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.

APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya

APPLICANT: Chalita-Eld, Pia M.
 APPLICANT: Paris, Mary

TITLE OF INVENTION: Nucleic acids and Corresponding Proteins

TITLE OF INVENTION: Encited 98P4B6 Useful in Treatment and Detection of Cancer

FILE REFERENCE: 51158-20016.24

CURRENT FILING DATE: 2003-06-04

PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned

PRIOR FILING DATE: 2002-12-20

PRIOR APPLICATION NUMBER: US60/370,387

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 19
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-19

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MESISMGSPKSLSTCTLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGHHVIGS 60
 1 MESISMGSPKSLSTCTLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGHHVIGS 60
 61 RNKFASEFFPHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRLHLLVGKILLIDVSNM 120
 61 RNKFASEFFPHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRLHLLVGKILLIDVSNM 120
 121 RINQVPSNAEYLAIFPDSLIVKGNVSAVALQGPDAAROYICNNIOAQOYIE 180
 121 RINQVPSNAEYLAIFPDSLIVKGNVSAVALQGPDAAROYICNNIOAQOYIE 180
 121 RINQVPSNAEYLAIFPDSLIVKGNVSAVALQGPDAAROYICNNIOAQOYIE 180
 181 LARQNLFIPIIDGLSSASAEIENLPLRLFTLRGPPVVAISLATFFLYSFVDVHHPA 240
 181 LARQNLFIPIIDGLSSASAEIENLPLRLFTLRGPPVVAISLATFFLYSFVDVHHPA 240
 241 RNOQSPFYKPIPIEVNKTPIVAITLLSLVYLAGLAAAYQYVGTGYRRFPFWLETWLQ 300
 241 RNOQSPFYKPIPIEVNKTPIVAITLLSLVYLAGLAAAYQYVGTGYRRFPFWLETWLQ 300
 241 RNOQSPFYKPIPIEVNKTPIVAITLLSLVYLAGLAAAYQYVGTGYRRFPFWLETWLQ 300
 301 CRKQGLSLFFPAMVAVASLCLPMRSEYFLNMAVQOVANINSNWEEVRIEMV 360
 301 CRKQGLSLFFPAMVAVASLCLPMRSEYFLNMAVQOVANINSNWEEVRIEMV 360
 361 ISFGMSGLSLAATVTSIPSVSNALNMRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
 361 ISFGMSGLSLAATVTSIPSVSNALNMRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
 421 EYYRFYTPPNFVLAIVLPSIVLIDLQCRYPD 454
 421 EYYRFYTPPNFVLAIVLPSIVLIDLQCRYPD 454

RESULT 6
 US-10-455-822-21
 Sequence 21, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Bid, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P486 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 21
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-21

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MESISMGSPKSLSTCTLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGHHVIGS 60
 1 MESISMGSPKSLSTCTLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGHHVIGS 60
 61 RNKFASEFFPHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRLHLLVGKILLIDVSNM 120
 61 RNKFASEFFPHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRLHLLVGKILLIDVSNM 120
 121 RINQVPSNAEYLAIFPDSLIVKGNVSAVALQGPDAAROYICNNIOAQOYIE 180
 121 RINQVPSNAEYLAIFPDSLIVKGNVSAVALQGPDAAROYICNNIOAQOYIE 180
 121 RINQVPSNAEYLAIFPDSLIVKGNVSAVALQGPDAAROYICNNIOAQOYIE 180
 181 LARQNLFIPIIDGLSSASAEIENLPLRLFTLRGPPVVAISLATFFLYSFVDVHHPA 240
 181 LARQNLFIPIIDGLSSASAEIENLPLRLFTLRGPPVVAISLATFFLYSFVDVHHPA 240
 241 RNOQSPFYKPIPIEVNKTPIVAITLLSLVYLAGLAAAYQYVGTGYRRFPFWLETWLQ 300
 241 RNOQSPFYKPIPIEVNKTPIVAITLLSLVYLAGLAAAYQYVGTGYRRFPFWLETWLQ 300
 241 RNOQSPFYKPIPIEVNKTPIVAITLLSLVYLAGLAAAYQYVGTGYRRFPFWLETWLQ 300
 301 CRKQGLSLFFPAMVAVASLCLPMRSEYFLNMAVQOVANINSNWEEVRIEMV 360
 301 CRKQGLSLFFPAMVAVASLCLPMRSEYFLNMAVQOVANINSNWEEVRIEMV 360
 361 ISFGMSGLSLAATVTSIPSVSNALNMRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
 361 ISFGMSGLSLAATVTSIPSVSNALNMRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
 421 EYYRFYTPPNFVLAIVLPSIVLIDLQCRYPD 454
 421 EYYRFYTPPNFVLAIVLPSIVLIDLQCRYPD 454

RESULT 7
 US-10-455-822-23
 Sequence 23, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Bid, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P486 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 23
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-23

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6,1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLENGINGIKDARKVTGVIGSGDPAKSLTIRLRCGYHVIGS 60
 DB 1 MESISMGSPKSLSETCLENGINGIKDARKVTGVIGSGDPAKSLTIRLRCGYHVIGS 60
 QY 61 RNPKPASEFPFHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
 DB 61 RNPKPASEFPFHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
 QY 121 RINQYPSNAEYIASLFPDLSLYKGNVVSAMALQGPDKASROYICSNNOIARQOYIE 180
 DB 121 RINQYPSNAEYIASLFPDLSLYKGNVVSAMALQGPDKASROYICSNNOIARQOYIE 180
 QY 181 LARQNLFPIDIGSLSSAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240
 DB 181 LARQNLFPIDIGSLSSAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240
 QY 241 RNOQSDFYKIPRIEIVNKTPIVAITLLSLVYLAGLLAAAYOLYGTGKRRFPFMLETWIQ 300
 DB 241 RNOQSDFYKIPRIEIVNKTPIVAITLLSLVYLAGLLAAAYOLYGTGKRRFPFMLETWIQ 300
 QY 301 CRKQGLISFFPAMVAVASLCIPMRSEHYFLNMAVQOVHANIENSNNEEVRIMY 360
 DB 301 CRKQGLISFFPAMVAVASLCIPMRSEHYFLNMAVQOVHANIENSNNEEVRIMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRBFSFIOSTLGYVALLISTFVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRBFSFIOSTLGYVALLISTFVLIYGMKRAFE 420
 QY 421 EYIRFPYTPPNFVALVLPISIVILDLLQCRYPD 454
 DB 421 EYIRFPYTPPNFVALVLPISIVILDLLQCRYPD 454

RESULT 8
 US-10-455-822-25

Sequence 25, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Mangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Eld, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 25
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-25

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6,1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLENGINGIKDARKVTGVIGSGDPAKSLTIRLRCGYHVIGS 60
 DB 1 MESISMGSPKSLSETCLENGINGIKDARKVTGVIGSGDPAKSLTIRLRCGYHVIGS 60
 QY 61 RNPKPASEFPFHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
 DB 61 RNPKPASEFPFHVVDVTHHEDALTNTNIIFVAIHREHYTSLMDLRHLVGLKILIDVSNM 120
 QY 121 RINQYPSNAEYIASLFPDLSLYKGNVVSAMALQGPDKASROYICSNNOIARQOYIE 180
 DB 121 RINQYPSNAEYIASLFPDLSLYKGNVVSAMALQGPDKASROYICSNNOIARQOYIE 180
 QY 181 LARQNLFPIDIGSLSSAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240
 DB 181 LARQNLFPIDIGSLSSAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240
 QY 241 RNOQSDFYKIPRIEIVNKTPIVAITLLSLVYLAGLLAAAYOLYGTGKRRFPFMLETWIQ 300
 DB 241 RNOQSDFYKIPRIEIVNKTPIVAITLLSLVYLAGLLAAAYOLYGTGKRRFPFMLETWIQ 300
 QY 301 CRKQGLISFFPAMVAVASLCIPMRSEHYFLNMAVQOVHANIENSNNEEVRIMY 360
 DB 301 CRKQGLISFFPAMVAVASLCIPMRSEHYFLNMAVQOVHANIENSNNEEVRIMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRBFSFIOSTLGYVALLISTFVLIYGMKRAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRBFSFIOSTLGYVALLISTFVLIYGMKRAFE 420
 QY 421 EYIRFPYTPPNFVALVLPISIVILDLLQCRYPD 454
 DB 421 EYIRFPYTPPNFVALVLPISIVILDLLQCRYPD 454

RESULT 9
 US-10-455-822-31

Sequence 31, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Mangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Eld, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98P4B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 31
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-31

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6,1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCLPENGINGIKDARKTVGVIGSGDPFAXSLTIRLRGCHVIGS 60
 DB 1 MESISMWSPKSLSETCLPENGINGIKDARKTVGVIGSGDPFAXSLTIRLRGCHVIGS 60
 QY 61 RNPFASEFPFPHVDVTHEDALTNTNIFVAIHREHTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPFASEFPFPHVDVTHEDALTNTNIFVAIHREHTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQPESSNAEYLAFLPDSLVKGFNVSAVALQGPKASROYICSNNIQARQVTE 180
 DB 121 RINQPESSNAEYLAFLPDSLVKGFNVSAVALQGPKASROYICSNNIQARQVTE 180
 QY 181 LARQNFIPIDIGSLSSAREIENLPLRFTLMRGVVAVALSLATFFFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDIGSLSSAREIENLPLRFTLMRGVVAVALSLATFFFLYSFVRDVIHPYA 240
 QY 241 RNOQSDFYKIPITVNTKPIVAITLLSLVYLAGLAAAYQLYGTKRFPFMLETMO 300
 DB 241 RNOQSDFYKIPITVNTKPIVAITLLSLVYLAGLAAAYQLYGTKRFPFMLETMO 300
 QY 301 CRKQGLSFFPAMVAVASLCIPMRRSERYFLNNAVQOVHANINENSWNEEVRIMY 360
 DB 301 CRKQGLSFFPAMVAVASLCIPMRRSERYFLNNAVQOVHANINENSWNEEVRIMY 360
 QY 361 ISFGIMSLGLSLAVALTSPVSANLNMREBFSFIQSTLGVALLISTFVLLIYGWRAFE 420
 DB 361 ISFGIMSLGLSLAVALTSPVSANLNMREBFSFIQSTLGVALLISTFVLLIYGWRAFE 420
 QY 421 EBYRFTPPNFVALVLPISIVILLDLQCRYPD 454
 DB 421 EBYRFTPPNFVALVLPISIVILLDLQCRYPD 454

RESULT 10
 US-10-455-822-33
 Sequence 33, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Mangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chalilte-Eld, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98PAB6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: NO. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 33
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-33

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6,1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCLPENGINGIKDARKTVGVIGSGDPFAXSLTIRLRGCHVIGS 60
 DB 1 MESISMWSPKSLSETCLPENGINGIKDARKTVGVIGSGDPFAXSLTIRLRGCHVIGS 60
 QY 61 RNPFASEFPFPHVDVTHEDALTNTNIFVAIHREHTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPFASEFPFPHVDVTHEDALTNTNIFVAIHREHTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQPESSNAEYLAFLPDSLVKGFNVSAVALQGPKASROYICSNNIQARQVTE 180
 DB 121 RINQPESSNAEYLAFLPDSLVKGFNVSAVALQGPKASROYICSNNIQARQVTE 180
 QY 181 LARQNFIPIDIGSLSSAREIENLPLRFTLMRGVVAVALSLATFFFLYSFVRDVIHPYA 240
 DB 181 LARQNFIPIDIGSLSSAREIENLPLRFTLMRGVVAVALSLATFFFLYSFVRDVIHPYA 240
 QY 241 RNOQSDFYKIPITVNTKPIVAITLLSLVYLAGLAAAYQLYGTKRFPFMLETMO 300
 DB 241 RNOQSDFYKIPITVNTKPIVAITLLSLVYLAGLAAAYQLYGTKRFPFMLETMO 300
 QY 301 CRKQGLSFFPAMVAVASLCIPMRRSERYFLNNAVQOVHANINENSWNEEVRIMY 360
 DB 301 CRKQGLSFFPAMVAVASLCIPMRRSERYFLNNAVQOVHANINENSWNEEVRIMY 360
 QY 361 ISFGIMSLGLSLAVALTSPVSANLNMREBFSFIQSTLGVALLISTFVLLIYGWRAFE 420
 DB 361 ISFGIMSLGLSLAVALTSPVSANLNMREBFSFIQSTLGVALLISTFVLLIYGWRAFE 420
 QY 421 EBYRFTPPNFVALVLPISIVILLDLQCRYPD 454
 DB 421 EBYRFTPPNFVALVLPISIVILLDLQCRYPD 454

RESULT 11
 US-10-455-822-35
 Sequence 35, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Mangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chalilte-Eld, Pia M.
 APPLICANT: Farris, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 98PAB6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: NO. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 35
LENGTH: 454
TYPE: PRF
ORGANISM: Homo sapien
US-10-455-822-35

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCIPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVYIGS 60
DB 1 MESISMGSPKSLSETCIPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVYIGS 60
QY 61 RNPKEASEFPFHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
DB 61 RNPKEASEFPFHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQPEESNAEYLAFLPDSLIYKGFNVASAMALQIGPKDSROYVICSNNIOARQOYIE 180
DB 121 RINQPEESNAEYLAFLPDSLIYKGFNVASAMALQIGPKDSROYVICSNNIOARQOYIE 180
QY 121 RINQPEESNAEYLAFLPDSLIYKGFNVASAMALQIGPKDSROYVICSNNIOARQOYIE 180
DB 121 RINQPEESNAEYLAFLPDSLIYKGFNVASAMALQIGPKDSROYVICSNNIOARQOYIE 180
QY 181 LARQNFPIPIDGSLSSAREIENLPRLFTLMRGVVAISLAFPEFLYSFVRDVIHBYA 240
DB 181 LARQNFPIPIDGSLSSAREIENLPRLFTLMRGVVAISLAFPEFLYSFVRDVIHBYA 240
QY 241 RNOQSDPFYKPIEIVNKTLPVIAITLLSLVYLAGLLAAAYOLYGTGKRRPPLMETLQ 300
DB 241 RNOQSDPFYKPIEIVNKTLPVIAITLLSLVYLAGLLAAAYOLYGTGKRRPPLMETLQ 300
QY 301 CRKQGLSFFPAMVAVAYSLCLPMRSEERYLFLMAYQOVHAINENSMNEEYVRIEMY 360
DB 301 CRKQGLSFFPAMVAVAYSLCLPMRSEERYLFLMAYQOVHAINENSMNEEYVRIEMY 360
QY 361 ISFGIMSLGLSLAVTIPSVSNALNWRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTIPSVSNALNWRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFFYTPPNFVALVLPSTVIIDLLQLCRYPD 454
DB 421 EBYRFFYTPPNFVALVLPSTVIIDLLQLCRYPD 454

RESULT 12
US-10-455-822-37

Sequence 37, Application US/10455822
Publication No. US20040048798A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Raitano, Arthur B.
APPLICANT: Ge, Mangmao
APPLICANT: Jakobovits, Aya
APPLICANT: Chailita-Bid, Pia M.
TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
TITLE OF INVENTION: Entitled 98P486 Useful in Treatment and Detection of Cancer
FILE REFERENCE: 51158-20016.24
CURRENT APPLICATION NUMBER: US/10/455,822
PRIOR FILING DATE: 2003-06-04
PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US60/370,387
PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US60/296,656
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US60/091,183
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US60/087,520
PRIOR FILING DATE: 1998-06-01
NUMBER OF SEQ ID NOS: 197
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 37
LENGTH: 454
TYPE: PRF
ORGANISM: Homo sapien
US-10-455-822-37

Query Match 100.0%; Score 2351; DB 12; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.1e-222;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCIPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVYIGS 60
DB 1 MESISMGSPKSLSETCIPNGINGIKDARKVTGVIGSGDFAKSLTIRLRGCHVYIGS 60
QY 61 RNPKEASEFPFHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
DB 61 RNPKEASEFPFHVVDVTHHEDALTKNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQPEESNAEYLAFLPDSLIYKGFNVASAMALQIGPKDSROYVICSNNIOARQOYIE 180
DB 121 RINQPEESNAEYLAFLPDSLIYKGFNVASAMALQIGPKDSROYVICSNNIOARQOYIE 180
QY 121 RINQPEESNAEYLAFLPDSLIYKGFNVASAMALQIGPKDSROYVICSNNIOARQOYIE 180
DB 121 RINQPEESNAEYLAFLPDSLIYKGFNVASAMALQIGPKDSROYVICSNNIOARQOYIE 180
QY 181 LARQNFPIPIDGSLSSAREIENLPRLFTLMRGVVAISLAFPEFLYSFVRDVIHBYA 240
DB 181 LARQNFPIPIDGSLSSAREIENLPRLFTLMRGVVAISLAFPEFLYSFVRDVIHBYA 240
QY 241 RNOQSDPFYKPIEIVNKTLPVIAITLLSLVYLAGLLAAAYOLYGTGKRRPPLMETLQ 300
DB 241 RNOQSDPFYKPIEIVNKTLPVIAITLLSLVYLAGLLAAAYOLYGTGKRRPPLMETLQ 300
QY 301 CRKQGLSFFPAMVAVAYSLCLPMRSEERYLFLMAYQOVHAINENSMNEEYVRIEMY 360
DB 301 CRKQGLSFFPAMVAVAYSLCLPMRSEERYLFLMAYQOVHAINENSMNEEYVRIEMY 360
QY 361 ISFGIMSLGLSLAVTIPSVSNALNWRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAVTIPSVSNALNWRSEFSIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRFFYTPPNFVALVLPSTVIIDLLQLCRYPD 454
DB 421 EBYRFFYTPPNFVALVLPSTVIIDLLQLCRYPD 454

RESULT 13
US-10-455-822-39

Sequence 39, Application US/10455822
Publication No. US20040048798A1
GENERAL INFORMATION:
APPLICANT: Agensys, Inc.
APPLICANT: Raitano, Arthur B.
APPLICANT: Ge, Mangmao
APPLICANT: Jakobovits, Aya
APPLICANT: Chailita-Bid, Pia M.
TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
TITLE OF INVENTION: Entitled 98P486 Useful in Treatment and Detection of Cancer
FILE REFERENCE: 51158-20016.24
CURRENT APPLICATION NUMBER: US/10/455,822
PRIOR FILING DATE: 2003-06-04
PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US60/370,387
PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 39
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-39

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSTCLPNGINGIKDARKVTGVIGSGDPFASKLTIRLRGCHVIGS 60
 DB 1 MESISMGSPKSLSTCLPNGINGIKDARKVTGVIGSGDPFASKLTIRLRGCHVIGS 60
 QY 61 RNPKFASFFPHVDVTHEDALTKNTIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKFASFFPHVDVTHEDALTKNTIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPDKASROYICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPDKASROYICSNNIQARQVIE 180
 QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPDKASROYICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPDKASROYICSNNIQARQVIE 180
 QY 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLRGCVVAISLATEFFLYSPYVDYIHVA 240
 DB 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLRGCVVAISLATEFFLYSPYVDYIHVA 240
 QY 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLRGCVVAISLATEFFLYSPYVDYIHVA 240
 DB 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLRGCVVAISLATEFFLYSPYVDYIHVA 240
 QY 241 RNOQSDFYKIPLEIVNKTLPYVAITLISLVYAGLAAAYQYVGTXYRRPPELMTLQ 300
 DB 241 RNOQSDFYKIPLEIVNKTLPYVAITLISLVYAGLAAAYQYVGTXYRRPPELMTLQ 300
 QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLNMAVQVHANIENSNNEEVRRIEMV 360
 DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLNMAVQVHANIENSNNEEVRRIEMV 360
 QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLNMAVQVHANIENSNNEEVRRIEMV 360
 DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLNMAVQVHANIENSNNEEVRRIEMV 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTIGYVALLISTFHVLIYGMKAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTIGYVALLISTFHVLIYGMKAFE 420
 QY 421 EYRRFTTPPNFVALVLPISIVLIDLQCRYPD 454
 DB 421 EYRRFTTPPNFVALVLPISIVLIDLQCRYPD 454

RESULT 14 US-10-455-822-78

Sequence 78, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Eld, Pia M.
 APPLICANT: Fairs, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 984B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 78
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-78

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSTCLPNGINGIKDARKVTGVIGSGDPFASKLTIRLRGCHVIGS 60
 DB 1 MESISMGSPKSLSTCLPNGINGIKDARKVTGVIGSGDPFASKLTIRLRGCHVIGS 60
 QY 61 RNPKFASFFPHVDVTHEDALTKNTIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKFASFFPHVDVTHEDALTKNTIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPDKASROYICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPDKASROYICSNNIQARQVIE 180
 QY 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPDKASROYICSNNIQARQVIE 180
 DB 121 RINQYPSNAEYLAFLPDSLIVKGFNVSAALQGPDKASROYICSNNIQARQVIE 180
 QY 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLRGCVVAISLATEFFLYSPYVDYIHVA 240
 DB 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLRGCVVAISLATEFFLYSPYVDYIHVA 240
 QY 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLRGCVVAISLATEFFLYSPYVDYIHVA 240
 DB 181 LARQNFIFIDIGSLSSAREIENLPLRLFTLRGCVVAISLATEFFLYSPYVDYIHVA 240
 QY 241 RNOQSDFYKIPLEIVNKTLPYVAITLISLVYAGLAAAYQYVGTXYRRPPELMTLQ 300
 DB 241 RNOQSDFYKIPLEIVNKTLPYVAITLISLVYAGLAAAYQYVGTXYRRPPELMTLQ 300
 QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLNMAVQVHANIENSNNEEVRRIEMV 360
 DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLNMAVQVHANIENSNNEEVRRIEMV 360
 QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLNMAVQVHANIENSNNEEVRRIEMV 360
 DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLNMAVQVHANIENSNNEEVRRIEMV 360
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTIGYVALLISTFHVLIYGMKAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWRREFSIQSTIGYVALLISTFHVLIYGMKAFE 420
 QY 421 EYRRFTTPPNFVALVLPISIVLIDLQCRYPD 454
 DB 421 EYRRFTTPPNFVALVLPISIVLIDLQCRYPD 454

RESULT 15 US-10-455-822-121

Sequence 121, Application US/10455822
 Publication No. US20040048798A1
 GENERAL INFORMATION:
 APPLICANT: Agensys, Inc.
 APPLICANT: Raitano, Arthur B.
 APPLICANT: Ge, Wangmao
 APPLICANT: Jakobovits, Aya
 APPLICANT: Chailita-Eld, Pia M.
 APPLICANT: Fairs, Mary
 TITLE OF INVENTION: Nucleic acids and Corresponding Proteins
 TITLE OF INVENTION: Entitled 984B6 Useful in Treatment and Detection of Cancer
 FILE REFERENCE: 51158-20016.24
 CURRENT APPLICATION NUMBER: US/10/455,822
 CURRENT FILING DATE: 2003-06-04
 PRIOR APPLICATION NUMBER: No. US20040048798A1 Yet Assigned
 PRIOR FILING DATE: 2002-12-20
 PRIOR APPLICATION NUMBER: US60/370,387
 PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US60/317,840
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: US60/296,656
 PRIOR FILING DATE: 2001-06-06
 PRIOR APPLICATION NUMBER: US60/091,183
 PRIOR FILING DATE: 1998-06-30
 PRIOR APPLICATION NUMBER: US60/087,520
 PRIOR FILING DATE: 1998-06-01
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 121
 LENGTH: 454
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-455-822-121

Query Match 100.0%; Score 2351; DB 12; Length 454;
 Best Local Similarity 100.0%; Pred. No. 6.1e-222;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MESISMNGSPKSLSETCLPNGINGIKDARKYTVIGSGDPFASLTITRLIRCGYHVIGS	60
Db	1	MESISMNGSPKSLSETCLPNGINGIKDARKYTVIGSGDPFASLTITRLIRCGYHVIGS	60
Qy	61	RNPKFASFPVVDVTHEDALTKNTIIVAIHREHYTSLMDLRHLVKGKILIDVSNM	120
Db	61	RNPKFASFPVVDVTHEDALTKNTIIVAIHREHYTSLMDLRHLVKGKILIDVSNM	120
Qy	121	RINQPEENAEYLASLPDLSIVKGFNVSAWALQGPKASRQVYICSNNICARQVIE	180
Db	121	RINQPEENAEYLASLPDLSIVKGFNVSAWALQGPKASRQVYICSNNICARQVIE	180
Qy	181	LARQINFLPIDLGSISAREIENLPLRLFTLMRGPVVAISLATPPFLYSFVRDVIHPYA	240
Db	181	LARQINFLPIDLGSISAREIENLPLRLFTLMRGPVVAISLATPPFLYSFVRDVIHPYA	240
Qy	241	RNOOSDFYKIPLEIVNKTLPVATITLSLVYLAGLAAAYQLYGTYRRPFWLFTMLQ	300
Db	241	RNOOSDFYKIPLEIVNKTLPVATITLSLVYLAGLAAAYQLYGTYRRPFWLFTMLQ	300
Qy	301	CRKQGLISFFPAMVHVAYSCLPMRSEERYLFLMAYQVHANIEENSWNEEYWRIMY	360
Db	301	CRKQGLISFFPAMVHVAYSCLPMRSEERYLFLMAYQVHANIEENSWNEEYWRIMY	360
Qy	361	ISFGIMSLGLSLAVTISPSVSNALNWRSEFIOSTLGYVALLISTFHVLIYGMKRAFE	420
Db	361	ISFGIMSLGLSLAVTISPSVSNALNWRSEFIOSTLGYVALLISTFHVLIYGMKRAFE	420
Qy	421	EYYRFTYTPNPFVLAIVLPSIYILDLQLCRYPD	454
Db	421	EYYRFTYTPNPFVLAIVLPSIYILDLQLCRYPD	454

Search completed: April 16, 2004, 12:46:36
 Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 12:38:02 ; Search time 22 Seconds

(without alignments)
1065.372 Million cell updates/sec

Title: US-09-455-486-6

Sequence: 1 MESSMMSGSPKSLSETCLPN.....ALVPSIVILDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 369414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 369414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A COMB pep:*
2: /cgn2_6/prodata/2/1aa/5A COMB pep:*
3: /cgn2_6/prodata/2/1aa/6A COMB pep:*
4: /cgn2_6/prodata/2/1aa/6A COMB pep:*
5: /cgn2_6/prodata/2/1aa/6A COMB pep:*
6: /cgn2_6/prodata/2/1aa/6A COMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	901	38.3	173	US-09-323-873A-8	Sequence 8, Appl
2	736	31.3	141	US-09-083-521-1	Sequence 1, Appl
3	717	30.5	339	US-09-323-873A-2	Sequence 2, Appl
4	717	30.5	339	US-09-685-166A-879	Sequence 879, App
5	144.5	6.1	227	US-09-655-270A-15	Sequence 15, Appl
6	144.5	6.1	227	US-09-651-941-17	Sequence 17, Appl
7	144.5	6.1	227	US-09-955-597-17	Sequence 17, Appl
8	107.5	4.6	695	US-08-487-886-2	Sequence 2, Appl
9	107.5	4.6	695	US-08-482-855-2	Sequence 2, Appl
10	107.5	4.6	695	US-08-474-886-2	Sequence 2, Appl
11	107.5	4.6	34	US-09-323-873A-20	Sequence 20, Appl
12	102.5	4.4	940	US-09-328-352-6165	Sequence 8165, Ap
13	101	4.3	476	US-09-316-083-3	Sequence 3, Appl
14	101	4.3	476	US-09-933-700-3	Sequence 3, Appl
15	101	4.3	724	US-09-252-991A-21494	Sequence 21494, A
16	99	4.2	365	US-09-170-486D-118	Sequence 118, App
17	99	4.2	365	US-09-170-486D-226	Sequence 226, App
18	99	4.2	365	US-09-364-425B-27	Sequence 27, Appl
19	99	4.2	531	US-08-724-974A-2	Sequence 2, Appl
20	97.5	4.1	692	US-08-460-576-2	Sequence 2, Appl
21	97	4.1	692	US-07-757-342D-6	Sequence 6, Appl
22	97	4.1	692	US-09-461-657B-6	Sequence 6, Appl
23	96	4.1	299	US-09-903-456-61	Sequence 61, Appl
24	96	4.1	1309	US-09-975-413A-10	Sequence 10, Appl
25	94.5	4.0	288	US-09-107-532A-6009	Sequence 6009, Ap
26	94.5	4.0	296	US-09-540-236-2922	Sequence 2922, Ap
27	94.5	4.0	345	US-09-489-039A-10740	Sequence 10740, A

28	94	4.0	284	US-09-903-456-83	Sequence 83, Appl
29	94	4.0	292	US-09-903-456-35	Sequence 35, Appl
30	94	4.0	293	US-09-903-456-45	Sequence 45, Appl
31	93	4.0	365	US-08-724-974A-3	Sequence 3, Appl
32	93	4.0	498	US-09-107-532A-7077	Sequence 7077, Ap
33	92.5	3.9	405	US-09-489-039A-9411	Sequence 9411, Ap
34	91.5	3.9	405	US-09-328-352-5605	Sequence 5605, Ap
35	91.5	3.9	495	US-08-985-343-2	Sequence 2, Appl
36	91	3.9	211	US-08-252-991A-23822	Sequence 23822, A
37	91	3.9	440	US-09-634-238-281	Sequence 281, App
38	91	3.9	550	US-09-543-681A-6026	Sequence 6026, Ap
39	91	3.9	808	US-09-114-001C-3105	Sequence 3105, Ap
40	90.5	3.8	340	US-09-543-681A-7503	Sequence 7503, Ap
41	90.5	3.8	388	US-09-107-532A-6381	Sequence 6381, Ap
42	90	3.8	500	US-08-178-093B-26	Sequence 26, Appl
43	90	3.8	970	US-08-795-927-7	Sequence 7, Appl
44	89.5	3.8	495	US-07-841-997A-2	Sequence 2, Appl
45	89.5	3.8	495	US-08-290-301-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-323-873A-8
Sequence 8, Application US/09323873A
Patent No. 6329503
GENERAL INFORMATION:
APPLICANT: Daniel E. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Kahan Leong
APPLICANT: Arthur B. Rattano
APPLICANT: Douglas C. Saffran
APPLICANT: Steve Chappell Mitchell
TITLE OF INVENTION: NOVEL SEQUENTINE TRANSMEMBRANE ANTIGENS
FILE REFERENCE: 128,160S02
CURRENT APPLICATION NUMBER: US/09/323, 873A
CURRENT FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087,520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 173
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-323-873A-8
Query Match 38.3%; Score 901; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.4e-86;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 246 DYKXPIPIRVNKLIVATITLSTLYLAGLLAAAYOLYGRKRRFPMTETWLQCRQL 305
DB 1 DYKPIPIRVNKLIVATITLSTLYLAGLLAAAYOLYGRKRRFPMTETWLQCRQL 60
QY 306 GLISFFPAMVAVYLCIPMRSEERYLLNAYQOVHANINENSWNEEVRIMETISGCI 365
DB 61 GLISFFPAMVAVYLCIPMRSEERYLLNAYQOVHANINENSWNEEVRIMETISGCI 120
QY 366 MSIGLSTLAATVSIPTSVSNALNWRBESFIOSTLGYVALLISTFHYLIIGMKRA 418
DB 121 MSIGLSTLAATVSIPTSVSNALNWRBESFIOSTLGYVALLISTFHYLIIGMKRA 173
RESULT 2
US-09-083-521-1
Sequence 1, Application US/09083521
Patent No. 6048970
GENERAL INFORMATION:

```
APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,521
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0527 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTUT10
CLONE: 1691243
US-09-083-521-1

Query Match
Best Local Similarity 31.3%; Score 736; DB 3; Length 141;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 MVAHVAISICLPMSRERYFLINMAYQVYHANIENSNMEEVWRLEMTISGIMSLGLSL 373
DB 1 MVAHVAISICLPMSRERYFLINMAYQVYHANIENSNMEEVWRLEMTISGIMSLGLSL 60
QY 374 LAVNISPSVSNALNMRESFIOSTLGYVALLISTFHVLIYGMKRAFESEYRFTPNFV 433
DB 61 LAVNISPSVSNALNMRESFIOSTLGYVALLISTFHVLIYGMKRAFESEYRFTPNFV 120
QY 434 LAVLPSIVILDLQLCRYPD 454
DB 121 LAVLPSIVILDLQLCRYPD 141

RESULT 3
US-09-323-873A-2
Sequence 2, Application US/09323873A
Patent No. 6329503
GENERAL INFORMATION:
APPLICANT: Daniel E. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Kahan Leong
APPLICANT: Arthur B. Saitano
APPLICANT: Douglas C. Saitano
APPLICANT: Steve Chappell Mitchell
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
FILE REFERENCE: 129 16TSU2
CURRENT APPLICATION NUMBER: US/09/323,873A
CURRENT FILING DATE: 1993-06-01
PRIOR APPLICATION NUMBER: 60/087,520
```

```
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 339
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-323-873A-2

Query Match
Best Local Similarity 30.5%; Score 717; DB 4; Length 339;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLMRGVVAVALATPFPLYSFVRDHPHRYANQOSDFYKPIETVNTLPTVAITLL 267
DB 67 LFPQNHPIKIAIATASITFTYTLREVIHPLATSHQGYFKIDIVINKVLPWVSITLL 126
QY 268 SLVYLAGLAAAYQVYGTGTRRPPWLEWLOCRKQGLISFFPAMVHVAISICLPMR 327
DB 127 ALVYLPVIAIAIVQLHNGTKYKKEPWLDKWMLTRKQFGLISFFFAVLHVAISISYPMR 186
QY 328 SRRYFLMVAQVYHANIENSNMEEVWRLEMTISGIMSLGLSLAVTSIPSVNALN 387
DB 187 SRYFLMVAQVYQVQNKEDAMIEHDVWRKEIYSLGIVGLAIALAVTSIPSVSGLT 246
QY 388 WREFSFIOSTLGYVALLISTFHVLIYGMKRAFESEYRFTPNFVALVLPISVIL 444
DB 247 WREFHYISKGIYVLLGTIHALIFANWKWIDIKQFVWYPPPTFMIAVFLPIVLI 303

RESULT 4
US-09-685-166A-879
Sequence 879, Application US/09685166A
Patent No. 6630305
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqun
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Scolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C21
CURRENT APPLICATION NUMBER: US/09/685,166A
CURRENT FILING DATE: 2000-10-10
NUMBER OF SEQ ID NOS: 898
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 879
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
US-09-685-166A-879

Query Match
Best Local Similarity 30.5%; Score 717; DB 4; Length 339;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLMRGVVAVALATPFPLYSFVRDHPHRYANQOSDFYKPIETVNTLPTVAITLL 267
```

```

DB 67 LPEGMHPIKIAIATLTLTYTLREVIHPLATSHOQPYKPIPLVINKVLPMSITLL 126
QY 268 SLVYLAGLAAAYQLYYGTIKRREPPWLEMTLOCRKQGLISFFPAMVAVSLCPMR 327
DB 127 ALVYLPVIAIAYQLHNGTKYKKEPFWLMDKXMLTRKQGLISFFPAMVAVSLCPMR 186
QY 328 SERFLNMAVQOQHANIENSMNEEYWRIRIEMTISFGIMSLGLSLAATISPSVNA 387
DB 187 SYRKLLNMAVQOQKREDAMIEDVMEIYVSLIAGLAIATLAVTISPSVDSLT 246
QY 388 WRESEFIOSTLGYVALISTFHVLIYGMKRAFESEYRYFTPPFVATLVPSIVIL 444
DB 247 WREHYIOSKIGTISLLGTITHALIFAMNKMIDIKQFVWTPPFPMVAVLPVILI 303

```

```

RESULT 5
US-09-655-270A-15
; Sequence 15, Application US/09655270A
; Patent No. 6329151
; GENERAL INFORMATION:
; APPLICANT: ROUVIERE, PIERRE E.
; TITLE OF INVENTION: High Density Sampling of Differentially Expressed Prokaryotic mR
; FILE REFERENCE: BC1011 US NA
; CURRENT APPLICATION NUMBER: US/09/655,270A
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/130,702
; PRIOR FILING DATE: 1999-February-19
; PRIOR APPLICATION NUMBER: 60/152,542
; PRIOR FILING DATE: 1999-September-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-655-270A-15

```

```

Query Match 6.1%; Score 144.5; DB 4; Length 227;
Best Local Similarity 28.0%; Pred. No. 6.4e-07;
Matches 60; Conservative 33; Mismatches 88; Indels 33; Gaps 9;
QY 25 IKDARKTVGVIGSGDFPAKSLTIRIRCGYHVIGSRNPKFASFPFH-----VVD 75
DB 1 MKSSKIAYVG--GTGPGKGLAYRFAAGMPVIGSRSAERAEBALEVRRAAGDGA 58
QY 76 VTHHEDALTKNIIFVAI---HREHYTSLMDLRLHLVSKILL-----DVSNNMRIN 123
DB 59 AADNAAADCPRIILLVVYPDGHRELVS---ELAPIFAGKLVSCVNPFGDKSGAYGLD 115
QY 124 QYPSNAEYIASLFPDSLIVKGFNVSA---WALQGFKDSRQYVIGSNNTQARQOYIE 180
DB 116 VEEGSAAEQLRDLVPATVVAAFHHLISAVNLMHEH-GP--LPEDVLVCGDDRSADDEVAR 172
QY 181 LABQLNFIPI-IDGLSLSAREIENLPLRLFTLMR 213
DB 173 LAVAITGRPGIDGALRVAROLEPLTAVLINVR 206

```

```

RESULT 6
US-09-651-941-17
; Sequence 17, Application US/09651941
; Patent No. 6355470
; GENERAL INFORMATION:
; APPLICANT: ROUVIER, PIERRE E
; APPLICANT: ROUVIER, PIERRE E
; APPLICANT: WALTERS, DANA M
; APPLICANT: RAINIER, RUSS
; TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
; FILE REFERENCE: BC1022 US NA
; CURRENT APPLICATION NUMBER: US/09/651,941
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,545
; PRIOR FILING DATE: 1999-10-03
; NUMBER OF SEQ ID NOS: 28

```

```

; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-651-941-17

```

```

Query Match 6.1%; Score 144.5; DB 4; Length 227;
Best Local Similarity 28.0%; Pred. No. 6.4e-07;
Matches 60; Conservative 33; Mismatches 88; Indels 33; Gaps 9;
QY 25 IKDARKTVGVIGSGDFPAKSLTIRIRCGYHVIGSRNPKFASFPFH-----VVD 75
DB 1 MKSSKIAYVG--GTGPGKGLAYRFAAGMPVIGSRSAERAEBALEVRRAAGDGA 58
QY 76 VTHHEDALTKNIIFVAI---HREHYTSLMDLRLHLVSKILL-----DVSNNMRIN 123
DB 59 AADNAAADCPRIILLVVYPDGHRELVS---ELAPIFAGKLVSCVNPFGDKSGAYGLD 115
QY 124 QYPSNAEYIASLFPDSLIVKGFNVSA---WALQGFKDSRQYVIGSNNTQARQOYIE 180
DB 116 VEEGSAAEQLRDLVPATVVAAFHHLISAVNLMHEH-GP--LPEDVLVCGDDRSADDEVAR 172
QY 181 LABQLNFIPI-IDGLSLSAREIENLPLRLFTLMR 213
DB 173 LAVAITGRPGIDGALRVAROLEPLTAVLINVR 206

```

```

RESULT 7
US-09-955-597-17
; Sequence 17, Application US/09955597
; Patent No. 6461856
; GENERAL INFORMATION:
; APPLICANT: ROUVIER, PIERRE E
; APPLICANT: WALTERS, DANA M
; APPLICANT: RAINIER, RUSS
; TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
; FILE REFERENCE: BC1022 US NA
; CURRENT APPLICATION NUMBER: US/09/955,597
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/152,545
; PRIOR FILING DATE: 1999-10-03
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-955-597-17

```

```

Query Match 6.1%; Score 144.5; DB 4; Length 227;
Best Local Similarity 28.0%; Pred. No. 6.4e-07;
Matches 60; Conservative 33; Mismatches 88; Indels 33; Gaps 9;
QY 25 IKDARKTVGVIGSGDFPAKSLTIRIRCGYHVIGSRNPKFASFPFH-----VVD 75
DB 1 MKSSKIAYVG--GTGPGKGLAYRFAAGMPVIGSRSAERAEBALEVRRAAGDGA 58
QY 76 VTHHEDALTKNIIFVAI---HREHYTSLMDLRLHLVSKILL-----DVSNNMRIN 123
DB 59 AADNAAADCPRIILLVVYPDGHRELVS---ELAPIFAGKLVSCVNPFGDKSGAYGLD 115
QY 124 QYPSNAEYIASLFPDSLIVKGFNVSA---WALQGFKDSRQYVIGSNNTQARQOYIE 180
DB 116 VEEGSAAEQLRDLVPATVVAAFHHLISAVNLMHEH-GP--LPEDVLVCGDDRSADDEVAR 172
QY 181 LABQLNFIPI-IDGLSLSAREIENLPLRLFTLMR 213
DB 173 LAVAITGRPGIDGALRVAROLEPLTAVLINVR 206

```

Sequence 2, Application US/06487886
 Patent No. 5744448
 GENERAL INFORMATION:
 APPLICANT: Kelton, Christie Ann
 APPLICANT: Schweichardt, Rene Lynn
 APPLICANT: Cheng, Shirley Van Yen
 APPLICANT: Nugent, No. 57444488en Patrice
 TITLE OF INVENTION: Human Follicle Stimulating
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Stephan P. Williams,
 ADDRESSEE: Ares-serono, Inc.
 STREET: Exchange Place, 37th floor
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
 COMPUTER: IBM PS/2, model 55 SX
 OPERATING SYSTEM: MS-DOS version 4.0
 SOFTWARE: VAX/VMS Massai via Kermit to IBM MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487,886
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/670,095
 FILING DATE: 15-MAR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Williams, Stephan P.
 REGISTRATION NUMBER: 28546
 REFERENCE/DOCKET NUMBER: US/252
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 723-1300
 TELEFAX: (617) 723-8923
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 695
 TYPE: Amino acid
 TOPOLOGY: Linear
 MOLECULE TYPE: Protein
 FEATURE:
 NAME/KEY: signal sequence
 LOCATION: -17 to -1
 IDENTIFICATION METHOD: hydrophobic
 FEATURE:
 NAME/KEY: putative amino-terminal extracellular domain
 LOCATION: 1 to 349
 IDENTIFICATION METHOD: similarity with other
 IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
 IDENTIFICATION METHOD: domains, hydrophilic
 FEATURE:
 NAME/KEY: transmembrane domain
 LOCATION: 350 to 613
 IDENTIFICATION METHOD: similarity to other G
 IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
 FEATURE:
 NAME/KEY: putative transmembrane region I
 LOCATION: 350 to 370
 IDENTIFICATION METHOD: similarity to other G
 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
 FEATURE:
 NAME/KEY: putative transmembrane region II
 LOCATION: 382 to 404
 IDENTIFICATION METHOD: similarity to other G
 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
 FEATURE:
 NAME/KEY: putative transmembrane region III
 LOCATION: 427 to 448

IDENTIFICATION METHOD: similarity to other G
 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
 FEATURE:
 NAME/KEY: putative transmembrane region IV
 LOCATION: 469 to 491
 IDENTIFICATION METHOD: similarity to other G
 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
 FEATURE:
 NAME/KEY: putative transmembrane region V
 LOCATION: 512 to 533
 IDENTIFICATION METHOD: similarity to other G
 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
 FEATURE:
 NAME/KEY: putative transmembrane region VI
 LOCATION: 557 to 580
 IDENTIFICATION METHOD: similarity to other G
 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
 FEATURE:
 NAME/KEY: putative transmembrane region VII
 LOCATION: 592 to 613
 IDENTIFICATION METHOD: similarity to other G
 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
 FEATURE:
 NAME/KEY: putative carboxy-terminal intracellular
 LOCATION: 614 to 678
 US-08-487-886-2
 Query Match 4.6%; Score 107.5; DB 1; Length 695;
 Best Local Similarity 17.0%; Pred. No. 0.028;
 Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;
 QY 79 HEDALTKNTNIFPAHREHYTSIMDRRLNG-----KILIDVSNMR 121
 DB 98 HETRIKANNL-LYINPEAFQNLPLNQLYLLISNTGKHLPDVHKHSLQKVLIDQDNIN 156
 QY 122 INQYPSNAEYVLSLPPDSLIVKGPVNVSAWALQLQPPKASROYVICSNNIQAQOVIEL 181
 DB 157 IH-----TIERNSEVGLSESVILML-----NKGIGQIHNCA----- 189
 QY 182 ARQUNFLPIDLGLSGARIEINPLRLFTLMGPPVVAISLTFEYSGFVADVHPAR 241
 DB 190 ---FNGTQDELNLSDNNMLELPNDVFGASGPVILDISRTIRHSLPSYGLBNLKRA 246
 QY 242 NQSDFFKPIPIEIVNKTLLPIVAITLLSLVY-----PPLIETMLQCRKL 305
 DB 247 RSTYNLKKL-----TLKLVALMEASLTYPHCCAFANMRROISELHPICNKSILROE 300
 QY 272 -----LAGLLAAVQLVYGTKRR-----PPLIETMLQCRKL 305
 DB 301 VDWYQTRGORSSIAEDNESSYGRGPDMTYEDYDLCEVVDVTCSPEDANPCEDIN 360
 QY 306 G---LSPFPAMVVA-----YSLCPMRSEERYFLNMAVQOVHANIN 347
 DB 361 GYNILRVLWIFISLAIATGIIIVVILTSQYKLTVP-----RFLMCONIAFAD----- 408
 QY 348 SWNEEYVRIEMVTSFGIMSLGLSLAVTSISVSNAIWMRE-----FSFIQSTL 398
 DB 409 -----LCIGIYVLLIASVDIHTKSYHNVAILDMQAGADAGFTVASEL 455
 QY 399 G---YVALLISTFH-----VLIVGMRAPSEERYRYPTPNFV-LA 435
 DB 456 SVYTLTITLERMTTTHANQDLCKVQLRRAASVWVGMIFAPALALFFIFGISTMKYS 515
 QY 436 LVLPISIVILDLQL 449
 DB 516 ICLPMDIDSPLSQL 529

RESULT 9
US-08-482-855-2
Sequence 2, Application US/08482855
Patent No. 6121016
GENERAL INFORMATION:
APPLICANT: Kelton, Christie Ann
APPLICANT: Schweichardt, Rene Lynn
APPLICANT: Cheng, Shirley Vui Yen
APPLICANT: Nugent, No. 6121016en Patrice
TITLE OF INVENTION: Human Follicle Stimulating
TITLE OF INVENTION: Hormone Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stephan P. Williams,
ADDRESSEE: Ares-serono, Inc.
STREET: Exchange Place, 37th floor
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
COMPUTER: IBM PS/2, Model 55 SX
OPERATING SYSTEM: MS-DOS version 4.0
SOFTWARE: VAX/VMS Massll via Kermit to IBM MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,855
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/670,085
FILING DATE: 15-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Stephan P.
REGISTRATION NUMBER: 28546
REFERENCE/DOCKET NUMBER: US/252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 723-1300
TELEFAX: (617) 723-8923
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 695
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Protein
FEATURE:
NAME/KEY: signal sequence
LOCATION: -17 to -1
IDENTIFICATION METHOD: hydrophobic
FEATURE:
NAME/KEY: putative amino-terminal extracellular domain
LOCATION: 1 to 349
IDENTIFICATION METHOD: similarity with other
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
IDENTIFICATION METHOD: domains, hydrophilic
FEATURE:
NAME/KEY: transmembrane domain
LOCATION: 350 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
FEATURE:
NAME/KEY: putative transmembrane region I
LOCATION: 350 to 370
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region II
LOCATION: 382 to 404
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length

FEATURE:
NAME/KEY: putative transmembrane region III
LOCATION: 427 to 448
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region IV
LOCATION: 469 to 491
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region V
LOCATION: 512 to 533
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VI
LOCATION: 557 to 580
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: putative transmembrane region VII
LOCATION: 592 to 613
IDENTIFICATION METHOD: similarity to other G
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
FEATURE:
NAME/KEY: domain
LOCATION: 614 to 678
US-08-482-855-2

Query Match 4.6%; Score 107.5; DB 3; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.028;
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

QY 79 HEDALKTNIIPAHIREHYTSIMDRHLVNG-----KIDVSNMR 121
DB 98 HETRIKANNL-LYVPEAFQNPJLQILISNTGKHLPDVAKHSLQKVLDDQDNIN 156
QY 122 INQPSNAEYASLPDGLIVGFNVASAMALQGPQASQVYICNNIQAQCVLRL 181
DB 157 IH-----TIRNSFVGLSFESVILML-----NKGIOIHNCA----- 189
QY 182 AROALNPIPDGLSSARIEINLPLEFLTWGPPVVAISLATFFLYSFVRDVHPYAR 241
DB 190 ---FNGTQDELNLSDNNNLELPDVFAGAGPVLLDISRTRHSLPSYGLENLKURA 246
QY 242 NQSDPKIPIEIVNKL-PIVAITLLSLVY-----PPLWLTLCRKQL 305
DB 247 RSTYNLKKLP-----TLKLVALMEASLTYPSCCAFANWRQISELHPICNKSILROE 300
QY 272 -----LAGLLAAVQLVYGYTKRR-----PPLWLTLCRKQL 305
DB 301 VDWYQTRGQSSGLADNDSYSRGDFMTYTEDYVLCNEVDVTSFPRDANPCEDIN 360
QY 306 G---LISFFPAMVVA-----YSLCLPMRSERYFLANAYQOVHANIN 347
DB 361 GYNILRLVIFISLAITGNIIVLVITTSQYCLVTP-----RFLMGNLAFAAD----- 408
QY 348 SWEDEEYWRLEMTISRGINSGLSLAATVTSISVSANLNR-----PSFIQSTL 398
DB 409 -----LCIGIYLLIASVDIHTKSOYHVAIDWQAGADAAAGFTVASEL 455
QY 399 G---YVALLISTFH-----VLITGWKRAPEEYRFPNPNFV-LA 435
DB 456 SVYTLTALITERHNTITHAMQJDCXVQLRHAASVMMGMIFAPAAALPFIIGISSWVKS 515
QY 436 LVPSIVILLDLQ 449


```

Query Match 4.3%; Score 101; DB 3; Length 476;
Best Local Similarity 20.7%; Pred. No. 0.075;
Matches 52; Conservative 48; Mismatches 85; Indels 66; Gaps 12

QY      62  NPKPASEEPFHVVDVTHHEDALTKTNI---IFVAHREHYT-----SLMD 103
      25  |N|P|V|A|F|S|I|N|---|K|T|L|A|K|E|I|F|T|N|Y|N|K|Y|S|Y|K|I|N|Q|I|N|H|I|P|Y|N|Y|K|
104  LRLHLVGLKLLIDVSNMNRINQYPESSAEYIASLF--PDSLYKGFNVVNSAMALQGPCKDA 161
      303  I|N|K|L|P|K|I|M|D|I|K|N|Y|W|L|A|G|F|A|A|D|S|F|L|S|Y|N|P|K|D|T|L|F|K|M---|
Db      303  I|N|K|L|P|K|I|M|D|I|K|N|Y|W|L|A|G|F|A|A|D|S|F|L|S|Y|N|P|K|D|T|L|F|K|M---| 347

QY      162  SHQVYICSNNIQARCOVILBARLNTIPIDIGLSAREENLPLRLTLMRGPPVVAIS 221
      348  -P|S|Y|I|S|-O|V|E|R|K|E|L|I|Y|I|D|E---S|P|L-S|I|S|V|K|K|G|N|K|L|D|F|K|I|F|T|T|E|L|I|M|K|
Db      348  -P|S|Y|I|S|-O|V|E|R|K|E|L|I|Y|I|D|E---S|P|L-S|I|S|V|K|K|G|N|K|L|D|F|K|I|F|T|T|E|L|I|M|K| 400

QY      222  LATFFPLSFVRVDVIPHVYANQOSDEPKIP---IEIVKTLPIVAITLLSLVYLAGLLA 277
      401  -----F|I|Y|F--D|K|E|P|L|H|D|N|K|Q|F|N|I|K|E|F|N|T|F|I|S|Y|W|N|N|R|V|F|G|I|V|L|S|E--Y|I|N|I|K|I|
Db      401  -----F|I|Y|F--D|K|E|P|L|H|D|N|K|Q|F|N|I|K|E|F|N|T|F|I|S|Y|W|N|N|R|V|F|G|I|V|L|S|E--Y|I|N|I|K|I| 451

QY      278  AAYLYGYGTYK 268
      452  D|N|D|Y|Y|Y|N|Y|N|K| 462
Db      452  D|N|D|Y|Y|Y|N|Y|N|K| 462

```

Query Match	4.3%;	Score 101;	DB 4;	Length 724;
Best Local Similarity	20.5%;	Pred. NO. 0.14;		
Matches 72;	Conservative 57;	Mismatches 105;	Indels 118;	Gaps 16;

```

QY 132 YLASLFPDLSLVKGFNVYSAMALQGPDKASROYVICSNNIOA-----ROQVIELARQUN 186
Db 298 YLSVYLPDLSL-SKTLAMWMAVIVCGTLFSA---LCVISLSLISGPHROPALDILRQA 352
QY 187 FIPIDL-GSLSAREIENLPRLFTLRGPVVAISLA-----TFEFLYSFVRDVIHP 238
Db 353 FRPIMLIGSLALAGVADHPRLIAGLGHTSICSTILANASALFTALFVRRFRPIAH- 411
QY 239 YARNQO-----SDFYKIPi-----EIVNK 257
Db 412 LIRNQPILERRLKRRSLHDVLQVLSLMEVPLVLVGISLFATFVSAGDSSSALRLALVCA 471
QY 258 TLPIVAITLLSLVYLAGLLAAAYQVYGTXYRRFPFPMLETLQCRKQGLISFFPAMVHV 317
Db 472 VLAIVAMTVIGLI-----RRRSRVAGPRRSAPYIEQ-----LOSFGYTLIHI 515
QY 318 AYSICLPYRRSERYLFLNMAYQOVHANIENSMNEEVRRIEM-YISFGIM----- 366
Db 516 FVVL-----FPIEVALRWGMSLIR-YAEGEGEIQISMKVVSFGTTLVAMLIWIL 564
QY 367 -----SLGL-----LSLAVTSIPSVSNALNMRPFGFIQSTLGYYALLIS 406
Db 565 TDTALQHSLSLGGKSRPNTRALTMPLIRNVL-----PATTAVIALIYA 608

```

Search completed: April 16, 2004, 12:41:47
 Job time : 24 secs

XX 06-DEC-2000; 2000WC-US033040.
 XX 06-DEC-1999; 99US-00455486.
 XX (UROG-) UROGENESYS INC.
 XX Afar DEH, Hubert RS, Raitano AB, Safiran DC, Mitchell SC;
 PI Faris M, Jakobovits A;
 XX WPI; 2001-367804/38.
 DR N-PSDB; AAD07072.
 XX
 PT New STEAP (six transmembrane epithelial antigen of the prostate)
 PT proteins, expressed in human cancers, useful for detecting and treating
 PT cancer.
 XX
 PS Claim 1; Fig 9A-9C; 187p; English.
 XX
 CC The present sequence is human six transmembrane epithelial antigen of the
 CC prostate (STEAP)-2 protein. STEAP is a member of cell surface serpentine
 CC transmembrane antigens. STEAP-2 gene is located on chromosome 7q21 and is
 CC used in gene therapy. Inhibiting the development or progression of a
 CC cancer (eg. prostate, colon, bladder, lung, ovarian and pancreatic)
 CC expressing STEAP or inhibiting growth or killing cells expressing STEAP
 CC in a patient, comprises administering a vaccine composition to the
 CC patient. Treating a patient with a cancer that expresses STEAP, or
 CC inhibiting growth or killing cells expressing STEAP, comprises
 CC administering to the patient a vector encoding single chain monoclonal
 CC antibody that comprises the variable domains of the heavy and light
 CC chains of the monoclonal antibody that specifically binds to STEAP, such
 CC that the vector delivers the single chain monoclonal antibody coding
 CC sequence to the cancer cells and the encoded single chain monoclonal
 CC antibody is expressed intracellularly. Note: This sequence is stated to
 CC be the same as that being shown as SEQ ID NO:8 (AA02841) in sequence
 CC listing of the specification. However both the sequences differ at
 CC several positions
 CC
 CC Sequence 454 AA;
 XX
 SQ

Query Match 100.0%; Score 2351; DB 4; Length 454;
 Best Local Similarity 100.0%; Pred. No. 5e-245;

Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCLPVGINGIKDARKVTGVIGSGDPFASKLITRLIRGCVHVVIGS 60
 DB 1 MESISMWSPKSLSETCLPVGINGIKDARKVTGVIGSGDPFASKLITRLIRGCVHVVIGS 60
 QY 61 RNPKFASEFPFHVVDVTHEDALTKNIIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKFASEFPFHVVDVTHEDALTKNIIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 QY 121 RINQPEENAEYLAFLPDSILVKGFNVSAMALQFGKDSRQYIICSNNOIARQVIE 180
 DB 121 RINQPEENAEYLAFLPDSILVKGFNVSAMALQFGKDSRQYIICSNNOIARQVIE 180
 QY 181 LARQNFPIPIDIGSSAREINLPLRFTLMRGVVAISLAPFFIYSGVRVVIHYA 240
 DB 181 LARQNFPIPIDIGSSAREINLPLRFTLMRGVVAISLAPFFIYSGVRVVIHYA 240
 QY 241 RNOQSDFFKIPFIEYNKTLPIVAITLLSVLAGLAAAYOLYYGKTRRPPMLWTLO 300
 DB 241 RNOQSDFFKIPFIEYNKTLPIVAITLLSVLAGLAAAYOLYYGKTRRPPMLWTLO 300
 QY 301 CRKQGLISFFPAAVAVASLCLPMKRSRRYFLMANVQOAHANENENMNEEVRIMY 360
 DB 301 CRKQGLISFFPAAVAVASLCLPMKRSRRYFLMANVQOAHANENENMNEEVRIMY 360
 QY 361 ISFGIMSLGLSLAVTSIPSVNMLNMEEFSTIGYVALLSTFHVLIYGMKFAFE 420
 DB 361 ISFGIMSLGLSLAVTSIPSVNMLNMEEFSTIGYVALLSTFHVLIYGMKFAFE 420
 QY 421 ESYRFPYTPNFVALVLPISVILLDLQICRYPD 454

DB 421 ESYRFPYTPNFVALVLPISVILLDLQICRYPD 454

RESULT 2
 ID AAU1018 standard; protein; 454 AA.
 XX
 AC AAU1018;

16-JAN-2002 (first entry)

Human ORF2 of Six-Transmembrane Protein of Prostate 1.

XX Human; Six-Transmembrane Protein of Prostate 1; STMP1; prostate cancer;
 XX benign prostatic hyperplasia; acute prostatitis; testicular cancer;
 KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
 KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
 KW liver cancer; lung cancer; cytostatic; ORF2.

XX Homo sapiens.

PN WO200172962-A2.

PD 04-OCT-2001.

PF 23-MAR-2001; 2001WC-US009410.

PR 24-MAR-2000; 2000US-0191929P.

PA (SAAT/) SAATCLOGLU F.

PI Saactioglu F;

XX WPI; 2001-662926/76.

DR N-PSDB; AAS15810, AAS15811.

PT New polynucleotide for the diagnosis, prevention and treatment for
 PT prostate and testis disorders, particularly prostate cancer, comprises
 PT prostate-specific or testis-specific nucleic acids.

PS Claim 1; Fig 4H; 114p; English.

CC The invention relates to substantially pure prostate-specific or testis-
 CC specific polypeptides and the nucleic acids encoding them. Also included
 CC are vectors and host cells expressing the proteins, a transgenic animal
 CC expressing the protein, antibodies against the proteins, probes for
 CC detecting the nucleic acids, antisense molecules for the nucleic acids
 CC and methods of isolating modulators of the proteins. Compounds that
 CC modulate the prostate specific or testis specific polypeptide are useful
 CC to diagnose, prevent or treat disorders of the testis or prostate
 CC particularly prostate cancer, benign prostatic hyperplasia, acute
 CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
 CC ascending or vanished testis. Other proliferative disorders for which the
 CC modulators may be used include lymphoma, leukaemia, melanoma, ovarian
 CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
 CC The present sequence is prostate specific protein, Six-Transmembrane
 CC protein of Prostate 1, STMP1, ORF2

XX Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 4; Length 454;
 Best Local Similarity 100.0%; Pred. No. 5e-245;

Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCLPVGINGIKDARKVTGVIGSGDPFASKLITRLIRGCVHVVIGS 60
 DB 1 MESISMWSPKSLSETCLPVGINGIKDARKVTGVIGSGDPFASKLITRLIRGCVHVVIGS 60
 QY 61 RNPKFASEFPFHVVDVTHEDALTKNIIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
 DB 61 RNPKFASEFPFHVVDVTHEDALTKNIIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120

QY 121 RINQPSNAEYLASLPDLSIVKGFNVSAWALQIGPKDASROYICSNNOARQOVI 180
 DB 121 RINQPSNAEYLASLPDLSIVKGFNVSAWALQIGPKDASROYICSNNOARQOVI 180
 QY 181 LARQNFIPIDIGLSASAREIENLPRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240
 DB 181 LARQNFIPIDIGLSASAREIENLPRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240
 QY 241 RNOQSDPYKPIPIEIVNKTLPVATLLSLVYLAAGLAAAYQLYGTYKTRFPPMLWTWQ 300
 DB 241 RNOQSDPYKPIPIEIVNKTLPVATLLSLVYLAAGLAAAYQLYGTYKTRFPPMLWTWQ 300
 QY 301 CRKQGLISLFFPFAVVAVSLCLPMRSEERYLFNMAVQOVHANIENSNNEEYVRLEMY 360
 DB 301 CRKQGLISLFFPFAVVAVSLCLPMRSEERYLFNMAVQOVHANIENSNNEEYVRLEMY 360
 QY 361 ISFGIMSLGILSLAATVTSIPSVSNALNMRSEFISQTLGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGILSLAATVTSIPSVSNALNMRSEFISQTLGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EBYRFPYTPPNFVLAIVLPSIVILDLQLCRYPD 454
 DB 421 EBYRFPYTPPNFVLAIVLPSIVILDLQLCRYPD 454

RESULT 3
 ABG61933 ID ABG61933 standard; protein; 454 AA.
 AC ABG61933;
 XX 15-AUG-2002 (first entry)
 DT 15-AUG-2002 (first entry)
 DE Prostate cancer-associated protein #134.
 XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
 OS Mammalia.
 PN WO200230268-A2.
 PD 18-APR-2002.
 PF 12-OCT-2001; 2001WO-US032045.
 XX 13-OCT-2000; 2000US-00687576.
 PR 08-DEC-2000; 2000US-00733288.
 PR 08-DEC-2000; 2000US-00733288.
 PR 24-JAN-2001; 2001US-0263957P.
 PR 16-MAR-2001; 2001US-0276791P.
 PR 16-MAR-2001; 2001US-0276888P.
 PR 06-APR-2001; 2001US-0281922P.
 PR 24-APR-2001; 2001US-0286214P.
 PR 30-APR-2001; 2001US-00847046.
 PR 04-MAY-2001; 2001US-0288589P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA
 PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
 DR WPI; 2002-471335/50.
 XX N-PSDB; ABK92252.
 PT Detecting a prostate cancer-associated transcript in a cell in a patient,
 PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
 PT by determining if prostate cancer-associated genes are expressed in a
 PT prostate tissue.
 XX
 PS Claim 27; Page 416; 436pp; English.
 CC The present invention relates to methods of detecting a prostate cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with prostate cancer-

CC associated polynucleotides (designated PC genes) that selectively
 CC hybridize to a sequence that is at least 80% identical to them. The
 CC prostate cancer-associated polynucleotide sequences are differentially
 CC expressed in prostate tumour tissue or in prostate cancer and are derived
 CC from the tissues of various organisms such as humans or other mammals
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
 CC associated genes are useful for diagnosing or treating prostate cancer,
 CC as well as for identifying modulators of prostate cancer or agents that
 CC inhibit prostate cancer. The nucleic acid sequences are particularly
 CC useful in gene therapy, as a vaccine or in antisense applications.
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins
 CC
 XX Sequence 454 AA:
 SQ
 Query Match 100.0%; Score 2351; DB 5; Length 454;
 Best Local Similarity 100.0%; Pred. No. 5e-245;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVYIGSGDFPKSLTRIRRCGVVVG 60
 DB 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVYIGSGDFPKSLTRIRRCGVVVG 60
 QY 61 RNPKFASFPFPHVVDVTHEDALTKNIIFFVAIHREHYTSLMDLHLLVGKILIDVSNM 120
 DB 61 RNPKFASFPFPHVVDVTHEDALTKNIIFFVAIHREHYTSLMDLHLLVGKILIDVSNM 120
 QY 121 RINQPSNAEYLASLPDLSIVKGFNVSAWALQIGPKDASROYICSNNOARQOVI 180
 DB 121 RINQPSNAEYLASLPDLSIVKGFNVSAWALQIGPKDASROYICSNNOARQOVI 180
 QY 181 LARQNFIPIDIGLSASAREIENLPRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240
 DB 181 LARQNFIPIDIGLSASAREIENLPRLFTLMRGPVVAISLATFFFLYSFVADVIHPYA 240
 QY 241 RNOQSDPYKPIPIEIVNKTLPVATLLSLVYLAAGLAAAYQLYGTYKTRFPPMLWTWQ 300
 DB 241 RNOQSDPYKPIPIEIVNKTLPVATLLSLVYLAAGLAAAYQLYGTYKTRFPPMLWTWQ 300
 QY 301 CRKQGLISLFFPFAVVAVSLCLPMRSEERYLFNMAVQOVHANIENSNNEEYVRLEMY 360
 DB 301 CRKQGLISLFFPFAVVAVSLCLPMRSEERYLFNMAVQOVHANIENSNNEEYVRLEMY 360
 QY 361 ISFGIMSLGILSLAATVTSIPSVSNALNMRSEFISQTLGYVALLISTFHVLIYGMKRAFE 420
 DB 361 ISFGIMSLGILSLAATVTSIPSVSNALNMRSEFISQTLGYVALLISTFHVLIYGMKRAFE 420
 QY 421 EBYRFPYTPPNFVLAIVLPSIVILDLQLCRYPD 454
 DB 421 EBYRFPYTPPNFVLAIVLPSIVILDLQLCRYPD 454

RESULT 4
 AAU76538 ID AAU76538 standard; protein; 454 AA.
 XX AAU76538;
 AC AAU76538;
 DT 05-JUN-2002 (first entry)
 DE Tumour-associated antigenic target protein, TAT138.
 XX
 XX TAT138, Tumour-associated Antigenic Target; tumour; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer;
 KW central nervous system cancer; liver cancer; bladder cancer; melanoma;
 KW pancreatic cancer; leukaemia; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 8..13
 FT Modified-site /note="N-myristoylation site"
 FT Modified-site 24..29

FI Hillan KJ, Watanabe CK, Wood WI, Yansura DG, Zhang Z;
 XX
 DR WPI: 2002-389270/41.
 DR N-PSDB: ABR50391.
 XX
 PT New polypeptide termed protein upregulated in metastatic prostate cancer
 PT and encoding polynucleotides, useful for identifying polypeptide
 PT antagonists for treating prostate cancer.
 XX
 PS Claim 23, Fig 2, 137pp; English.

The invention relates to an isolated human protein upregulated in metastatic prostate cancer (PUMPCn). PRO32303 polypeptide, a sequence 80% identical to PRO32303 and the sequence as encoded by cDNA insert of the vector deposited as ATCC Deposit No. PTA-2513 (DNA185171-2994) or 26/9/2000. Also included are the polynucleotide encoding the protein (or a DNA sequence 80% identical to the polynucleotide) and one that hybridises to complement of the polynucleotide), a vector comprising the polynucleotide, a polynucleotide deposited with ATCC under accession number PTA-2513 (DNA185171-2994), a host cell comprising the vector, preparation of PRO32303, a chemie molecule comprising PRO32303 fused to a heterologous amino acid sequence, an anti-PRO32303 antibody an agonist/antagonist of PRO32303, and diagnosing the presence of prostate cancer in a mammal by: (a) contacting a microarray diagnostic with a DNA185171-2994 probe, detecting and quantifying hybridisation of DNA185171-2994 probe in prostate cancer tissue compared with normal tissue and determining if DNA185171-2994 is overexpressed; or (b) contacting a tissue of the mammal with an anti-PRO32303 antibody and detecting the binding of the antibody to a component of the tissue, where binding is indicative of the presence of prostate cancer in the mammal. The antibody is useful for treating prostate cancer in mammal which is androgen independent prostate cancer, that has metastasised to another portion of the body, where the antibody is not conjugated with a cytotoxic agent and the method further comprises administering a chemotherapeutic agent to the mammal and for diagnosing the presence of prostate cancer in a mammal. The PRO32303 polynucleotide is useful for generating either transgenic animals or knock out animals which in turn are useful in the development and screening of therapeutically useful reagents or for use in gene therapy and for chromosome identification. PRO32303 and polynucleotide may also be used for tissue typing and PRO32303 may also be employed as a therapeutic agents and for screening compounds to identify those that mimic the PRO32303 polypeptide (agonists) or prevent the effect of the PRO32303 polypeptide (antagonists). PRO32303 is further useful for the affinity purification of PRO32303 from recombinant cell culture or natural sources. The present sequence represents PRO32303

SQ Sequence 454 AA;

Query Match	100.0%	Score 2351	DB 5	Length 454
Best Local Similarity	100.0%	Pred. No. 5e-245		
Matches 454, Conservative	0	Mismatches	0	Indels 0
				Gaps 0

Qy	1	MES1SMNGSKSLSECTCPGNGINGIDAKRVAVIGSGDPAKSITIRLLRFGYHVIGS	60
Db	1	MES1SMNGSKSLSECTCPGNGINGIDAKRVAVIGSGDPAKSITIRLLRFGYHVIGS	60
Qy	61	RNPFASFEPPHYVDVTHHEDALTKNIIIFVAIHEHNTSLAMDRLHVGKILLIDVENNM	120
Db	61	RNPFASFEPPHYVDVTHHEDALTKNIIIFVAIHEHNTSLAMDRLHVGKILLIDVENNM	120
Qy	121	RINQYPSNMEYLASLPDLSLYKGNVNSAVALDGPMDASROYICSNNOIQAQGVIE	180
Db	121	RINQYPSNMEYLASLPDLSLYKGNVNSAVALDGPMDASROYICSNNOIQAQGVIE	180
Qy	181	LAROLANTLPIDGSLGSAFIEMLPLRLTLMRGPVVAIISATFEFLYSVRVIVHPYA	240
Db	181	LAROLANTLPIDGSLGSAFIEMLPLRLTLMRGPVVAIISATFEFLYSVRVIVHPYA	240
Qy	241	RNQGSDPYKPIPIELVNKTLPIVAITLLSLVYLAGLIAAAVQLYGTGKRRRPPMLRWLQ	3000
Db	241	RNQGSDPYKPIPIELVNKTLPIVAITLLSLVYLAGLIAAAVQLYGTGKRRRPPMLRWLQ	3000

QY	301	CRKQGLISFFPAMVAVVSLCLPMRSEKYLFLNMAVQOVHANIENSMNEEVMRIEMY	361
Db	301	CRKQGLISFFPAMVAVVSLCLPMRSEKYLFLNMAVQOVHANIENSMNEEVMRIEMY	360
QY	361	ISFGIMSLGLSLAVTISIPSVSNALNWRFEFSIOSTLGYVALLISTFVLIYGMRAFE	420
Db	361	ISFGIMSLGLSLAVTISIPSVSNALNWRFEFSIOSTLGYVALLISTFVLIYGMRAFE	420
QY	421	EEYRREYTPPNFVALVLPSTVILDLQLCRYPD	454
Db	421	EEYRREYTPPNFVALVLPSTVILDLQLCRYPD	454
RESULT 6			
ABU08893			
ID	ABU08893	standard; protein; 454 AA.	
XX			
AC	ABU08893;		
DT	09-OCT-2003	(first entry)	
XX			
DB	Tumour-associated antigenic target 138 (TAT138).		
KW	Human; tumour-associated antigenic target polypeptide; TAT; tumour;		
XX	gene therapy; cytostatic; tissue typing; prostate tumour; cancer.		
OS	Homo sapiens.		
XX	US2003060612-A1.		
PN	27-MAR-2003.		
XX			
PF	22-JUN-2001; 2001US-00888257.		
XX			
PR	28-OCT-1997; 97US-0063540P.		
PR	17-JUN-1998; 98US-0089653P.		
PR	10-SEP-1998; 98US-0099792P.		
PR	08-OCT-1998; 98US-0103678P.		
PR	02-JUN-1999; 99WO-US012253.		
PR	01-SEP-1999; 99WO-US020111.		
PR	18-FEB-2000; 2000WO-US00434Z.		
PR	02-MAR-2000; 2000WO-US005843.		
PR	30-MAR-2000; 2000WO-US008439.		
PR	24-AUG-2000; 2000WO-US023328.		
PR	26-SEP-2000; 2000US-0235451P.		
PR	01-DEC-2000; 2000WO-US03678P.		
PR	28-FEB-2001; 2001WO-US006520.		
PR	01-MAR-2001; 2001WO-US006666.		
XX-			
PA	(GETH) GENENTECH INC.		
PI	Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V;		
PI	Wood WI, Wu TD, Zhang Z;		
XX			
DR	WPI: 2003-576355/54.		
XX	N-ESDB: ACD25893.		
PT	New nucleic acid, useful for the manufacture of a medicament for		
PT	diagnosing or treating tumor in a mammal.		
PS	Claim 12; Fig 10; 71pp; English.		
XX			
CC	The invention discloses human nucleic acids encoding tumour-associated		
CC	antigenic target (TAT) polypeptides, with or without their associated		
CC	signal peptide. Also disclosed is an antibody that specifically binds to		
CC	the TAT polypeptides, a method for detecting the presence of a tumour in		
CC	a mammal and a method for killing a cancer cell expressing the TAT		
CC	polypeptide. The nucleotide sequences are useful in preparing TAT		
CC	polypeptides by recombinant techniques and in gene therapy (e.g. for		
CC	replacement of defective gene). The TAT polypeptides are useful as		
CC	therapeutic agents and for detecting the presence, prevention and/or		
CC	treatment of a tumour, such as colon, breast or prostate tumour. The TAT		
CC	polypeptides and nucleic acids may also be used diagnostically for tissue		

CC Typing. The sequence presented is the TAT138 polypeptide of the invention
 XX Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 7; Length 454;
 Best Local Similarity 100.0%; Pred. No. 5e-245;
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCLPENGINGIKDARKVTVGVIGSGDFAKSLTIRLIRCGYHVIVGS 60
 DB 1 MESISMWSPKSLSETCLPENGINGIKDARKVTVGVIGSGDFAKSLTIRLIRCGYHVIVGS 60
 QY 61 RNPFFASEFPFPHVVDVTHHEDALTKNTIIFVAIHREHTSLMDLPHLVGKTLIDVSNM 120
 DB 61 RNPFFASEFPFPHVVDVTHHEDALTKNTIIFVAIHREHTSLMDLPHLVGKTLIDVSNM 120
 QY 121 RINQYPSNAEYTLASLPDLSLVKGFNVVSAMALQGPXASROYIICSNMIOARQOYIE 180
 DB 121 RINQYPSNAEYTLASLPDLSLVKGFNVVSAMALQGPXASROYIICSNMIOARQOYIE 180
 QY 181 LARQNLNFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240
 DB 181 LARQNLNFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240
 QY 241 RNOQSDFYKPIPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYGKYRRFPFMLETWLQ 300
 DB 241 RNOQSDFYKPIPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYGKYRRFPFMLETWLQ 300
 QY 301 CRKQGLISFFPFAVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENSNNEEVRIRIEM 360
 DB 301 CRKQGLISFFPFAVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENSNNEEVRIRIEM 360
 QY 361 ISFGIMSIGLSLAAVTSIPSVSNALNMWREFSIQSTLGYVALLISTFVLLIYGMKRAFE 420
 DB 361 ISFGIMSIGLSLAAVTSIPSVSNALNMWREFSIQSTLGYVALLISTFVLLIYGMKRAFE 420
 QY 421 EYYRFFYTPNPFVLAIVLPISIVILDLQLCRYPD 454
 DB 421 EYYRFFYTPNPFVLAIVLPISIVILDLQLCRYPD 454

RESULT 7
 ADB65001 standard; protein; 454 AA.

XX ADB65001;
 XX 04-DEC-2003 (first entry)
 XX Human protein encoded by clone PROST20168600.
 XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
 XX cell regeneration; membrane protein; signal transduction-related protein;
 XX transcription-related protein; osteoporosis; neurological disease;
 XX cancer; tumour.
 XX Homo sapiens.
 XX EP1308459-A2.
 XX 07-MAY-2003.
 XX 28-MAR-2002; 2002EP-00007401.
 XX 05-NOV-2001; 2001JP-00379298.
 XX 25-JAN-2002; 2002US-00350978.
 XX (HELI-) HELIX RES INST.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX Isogai T, Sugiyama T, Otsuki T, Makamatsu A, Sato H, Iehi S;
 XX Yamamoto U, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 XX Seki N, Yoshikawa T, Otsuka W, Nagahari K, Masuno Y;

XX WPI, 2003-450961/43.
 DR N-PSDB; ADB63031.
 XX New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.
 XX Claim 1; Page: 222pp; English.

The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or peptide
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesizing the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC medicines may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence data for this patent is not represented in the printed by the
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.

Sequence 454 AA;

Query Match 99.9%; Score 2348; DB 7; Length 454;
 Best Local Similarity 99.8%; Pred. No. 1.1e-244;
 Matches 453; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCLPENGINGIKDARKVTVGVIGSGDFAKSLTIRLIRCGYHVIVGS 60
 DB 1 MESISMWSPKSLSETCLPENGINGIKDARKVTVGVIGSGDFAKSLTIRLIRCGYHVIVGS 60
 QY 61 RNPFFASEFPFPHVVDVTHHEDALTKNTIIFVAIHREHTSLMDLPHLVGKTLIDVSNM 120
 DB 61 RNPFFASEFPFPHVVDVTHHEDALTKNTIIFVAIHREHTSLMDLPHLVGKTLIDVSNM 120
 QY 121 RINQYPSNAEYTLASLPDLSLVKGFNVVSAMALQGPXASROYIICSNMIOARQOYIE 180
 DB 121 RINQYPSNAEYTLASLPDLSLVKGFNVVSAMALQGPXASROYIICSNMIOARQOYIE 180
 QY 181 LARQNLNFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240
 DB 181 LARQNLNFIPIIDGLSSAREIENLPLRLFTLMRGPVVAISLATFFLYSFVRDVIHPYA 240
 QY 241 RNOQSDFYKPIPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYGKYRRFPFMLETWLQ 300
 DB 241 RNOQSDFYKPIPIEIVNKTLPPIVATITLSLVYLAGLLAAAYQLYGKYRRFPFMLETWLQ 300
 QY 301 CRKQGLISFFPFAVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENSNNEEVRIRIEM 360
 DB 301 CRKQGLISFFPFAVHVAVSLCLPMRSEERYFLNMAVYQOVHANIENSNNEEVRIRIEM 360
 QY 361 ISFGIMSIGLSLAAVTSIPSVSNALNMWREFSIQSTLGYVALLISTFVLLIYGMKRAFE 420
 DB 361 ISFGIMSIGLSLAAVTSIPSVSNALNMWREFSIQSTLGYVALLISTFVLLIYGMKRAFE 420
 QY 421 EYYRFFYTPNPFVLAIVLPISIVILDLQLCRYPD 454
 DB 421 EYYRFFYTPNPFVLAIVLPISIVILDLQLCRYPD 454

transmembrane protein differentially expressed in prostate and lung
tumour TMPL
Sequence 490 AA;

Query Match 97.6%; Score 2294; DB 7; Length 490;
Best Local Similarity 100.0%; Pred. No. 8.4e-239;
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSECTCPNGINGIKDARKVTGVYIGSDPFAKSLTITLRCGHVYIGS 60
DB 1 MESISMWSPKSLSECTCPNGINGIKDARKVTGVYIGSDPFAKSLTITLRCGHVYIGS 60
QY 61 RNPFASEFFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLMDRHLVGGKILIDVSNM 120
DB 61 RNPFASEFFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLMDRHLVGGKILIDVSNM 120
QY 121 RINQYPSNAEYLAFLPDSLVKGFNVASAMALQGPDKASROYVICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLAFLPDSLVKGFNVASAMALQGPDKASROYVICSNNIQARQVIE 180
QY 181 LARQNLFIPIIDGSLSSAREIENLPLRLFTLRGPVVAISLATFFLYSFVRDVIHPYA 240
DB 181 LARQNLFIPIIDGSLSSAREIENLPLRLFTLRGPVVAISLATFFLYSFVRDVIHPYA 240
QY 241 RNOQSDPYKPIEYINKTLPVAVITLSLVLAGLLAAVQLYGTYKRRPPLLEMTLQ 300
DB 241 RNOQSDPYKPIEYINKTLPVAVITLSLVLAGLLAAVQLYGTYKRRPPLLEMTLQ 300
QY 301 CRKQGLISFFPFAVHVAVSLCPMRSEERYFLNMAVQOVHANIENTSNEEVRRIEM 360
DB 301 CRKQGLISFFPFAVHVAVSLCPMRSEERYFLNMAVQOVHANIENTSNEEVRRIEM 360
QY 361 ISFGIMSLGSLSLAVTSIPSVSNALNMRPESFQSTLGYVALLISTFHLIYGWKRAFE 420
DB 361 ISFGIMSLGSLSLAVTSIPSVSNALNMRPESFQSTLGYVALLISTFHLIYGWKRAFE 420
QY 421 EEYRFFYTPPNFVALVLPISIVIL 444
DB 421 EEYRFFYTPPNFVALVLPISIVIL 444

RESULT 10
AAU0187
ID AAU0187 standard; protein; 490 AA.
XX AC AAU0187;
XX DT 16-JAN-2002 (first entry)
XX DE Human Six-Transmembrane Protein of Prostate 1, STMPL.
XX KM Human; Six-Transmembrane Protein of Prostate 1; STM; 1; prostate cancer;
XX KM benign prostatic hyperplasia; acute prostatic cancer;
XX KM cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
XX KM leukemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
XX KM liver cancer; lung cancer; cytostatic.
XX OS Homo sapiens.
XX FH Key
FH FT 208..230 Location/Qualifiers
FT /label= Transmembrane_domain_1
FT 255..273
FT /label= Transmembrane_domain_2
FT 304..325
FT /label= Transmembrane_domain_3
FT 360..380
FT /label= Transmembrane_domain_4
FT 393..415
FT /label= Transmembrane_domain_5
FT 432..452
FT /label= Transmembrane_domain_6

XX PN W0200172962-A2.
XX PD 04-OCT-2001.
XX PF 23-MAR-2001; 2001W0-US009410.
XX PR 24-MAR-2000; 2000US-0191929P.
XX PA (SAAT/) SAATCIOGLU F.
XX PT Saatioglu F;
XX DR WPI; 2001-662926/76.
XX DR N-PSDB; AAS15793, AAS15801, AAS15802.
XX PT New polynucleotide for the diagnosis, prevention and treatment for
XX PT prostate and testis disorders, particularly prostate cancer, comprises
XX PT prostate-specific or testis-specific nucleic acids.

XX PS Claim 1; Fig 4B; 114pp; English.

XX CC The invention relates to substantially pure prostate-specific or testis-
XX CC specific polypeptides and the nucleic acids encoding them. Also included
XX CC are vectors and host cells expressing the proteins, a transgenic animal
XX CC expressing the protein, antibodies against the proteins, probes for
XX CC detecting the nucleic acids, antisense molecules for the nucleic acids
XX CC and methods of isolating modulators of the proteins. Compounds that
XX CC modulate the prostate specific or testis specific polypeptide are useful
XX CC to diagnose, prevent or treat disorders of the testis or prostate
XX CC particularly prostate cancer, benign prostatic hyperplasia, acute
XX CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,
XX CC ascending or vanished testis. Other proliferative disorders for which the
XX CC modulators may be used include lymphoma, leukemia, melanoma, ovarian
XX CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
XX CC The present sequence represents a prostate specific protein, Six-
XX CC transmembrane protein of prostate 1, STM, 1

XX SQ Sequence 490 AA;

Query Match 97.4%; Score 2290; DB 4; Length 490;
Best Local Similarity 99.8%; Pred. No. 2.3e-238;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSECTCPNGINGIKDARKVTGVYIGSDPFAKSLTITLRCGHVYIGS 60
DB 1 MESISMWSPKSLSECTCPNGINGIKDARKVTGVYIGSDPFAKSLTITLRCGHVYIGS 60
QY 61 RNPFASEFFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLMDRHLVGGKILIDVSNM 120
DB 61 RNPFASEFFPHVVDVTHHEDALTKNTIIFVAIHREHYTSLMDRHLVGGKILIDVSNM 120
QY 121 RINQYPSNAEYLAFLPDSLVKGFNVASAMALQGPDKASROYVICSNNIQARQVIE 180
DB 121 RINQYPSNAEYLAFLPDSLVKGFNVASAMALQGPDKASROYVICSNNIQARQVIE 180
QY 181 LARQNLFIPIIDGSLSSAREIENLPLRLFTLRGPVVAISLATFFLYSFVRDVIHPYA 240
DB 181 LARQNLFIPIIDGSLSSAREIENLPLRLFTLRGPVVAISLATFFLYSFVRDVIHPYA 240
QY 241 RNOQSDPYKPIEYINKTLPVAVITLSLVLAGLLAAVQLYGTYKRRPPLLEMTLQ 300
DB 241 RNOQSDPYKPIEYINKTLPVAVITLSLVLAGLLAAVQLYGTYKRRPPLLEMTLQ 300
QY 301 CRKQGLISFFPFAVHVAVSLCPMRSEERYFLNMAVQOVHANIENTSNEEVRRIEM 360
DB 301 CRKQGLISFFPFAVHVAVSLCPMRSEERYFLNMAVQOVHANIENTSNEEVRRIEM 360
QY 361 ISFGIMSLGSLSLAVTSIPSVSNALNMRPESFQSTLGYVALLISTFHLIYGWKRAFE 420
DB 361 ISFGIMSLGSLSLAVTSIPSVSNALNMRPESFQSTLGYVALLISTFHLIYGWKRAFE 420
QY 421 EEYRFFYTPPNFVALVLPISIVIL 444

Db 421 EEYRPTPTNFTLALVPSIVL 444

RESULT 11
AAU10189 standard; protein; 419 AA.

XX AAU10189;
XX 16-JAN-2002 (first entry)
XX Human ORF3 of Six-Transmembrane Protein of Prostate 1.
XX Human; Six-Transmembrane Protein of Prostate 1; STM; 1; prostate cancer;
XX benign prostatic hyperplasia; acute prostatic; testicular cancer;
XX cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
XX leukemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
XX liver cancer; lung cancer; cytostatic; ORF3.
XX Homo sapiens.
XX WO200172962-A2.
XX 04-OCT-2001.
XX 23-MAR-2001; 2001WO-US009410.
XX 24-MAR-2000; 2000US-0191929P.
XX (SAAT/) SAATCIGLU F.
XX Saatioglu F;
XX WPI; 2001-662926/76.
XX N-PSDB; AAS15813.
XX New polynucleotide for the diagnosis, prevention and treatment for
XX prostate and testis disorders, particularly prostate cancer, comprises
XX prostate-specific or testis-specific nucleic acids.
XX Claim 1; Fig 4K; 114pp; English.
XX The invention relates to substantially pure prostate-specific or testis-
XX specific polypeptides and the nucleic acids encoding them. Also included
XX are vectors and host cells expressing the proteins, a transgenic animal
XX detecting the protein, antibodies against the proteins, probes for
XX detecting the nucleic acids, antisense molecules for the nucleic acids
XX and methods of isolating modulators of the proteins. Compounds that
XX modulate the prostate specific or testis specific polypeptide are useful
XX to diagnose, prevent or treat disorders of the testis or prostate
XX particularly prostate cancer, benign prostatic hyperplasia, acute
XX prostaticitis, testicular cancer, cryptorchidism, undescended, retractile,
XX ascending or vanished testis. Other proliferative disorders for which the
XX modulators may be used include lymphoma, leukemia, melanoma, ovarian
XX cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.
XX The present sequence is prostate specific protein, Six-Transmembrane
XX Protein of Prostate 1, STM,1, ORF3
XX Sequence 419 AA;
SQ Query Match 86.6%; Score 2036; DB 4; Length 419;
Best Local Similarity 99.7%; Pred. No. 5.5e-211;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSTCLPNGINGIKDKRYTVGVISGDPKSLTIRLRGCVHVVIGS 60
Db 1 MESISMWSPKSLSTCLPNGINGIKDKRYTVGVISGDPKSLTIRLRGCVHVVIGS 60
QY 61 RNPKEAFEPFHVVDVTHEDALTNTNIIIPVAIHEHYTSLMDIRHLVNGKITLIDVSNM 120
Db 61 RNPKEAFEPFHVVDVTHEDALTNTNIIIPVAIHEHYTSLMDIRHLVNGKITLIDVSNM 120

QY 121 RINQYPSNNAEYASLPFDSLIVKGFNVASAMALQLGPKDASQVITCSNNIQARQVIE 180
Db 121 RINQYPSNNAEYASLPFDSLIVKGFNVASAMALQLGPKDASQVITCSNNIQARQVIE 180
QY 181 LAROLNFTPIDIGSLSSAREIENPLRLFTLTMGPVVAISLATFFLYSFVSDVHIPYA 240
Db 181 LAROLNFTPIDIGSLSSAREIENPLRLFTLTMGPVVAISLATFFLYSFVSDVHIPYA 240
QY 241 RNOQSDPYKPIEIVNKTLPVATITLLSLVYLAGLLAAAYQLYGTYKRRFPPELFTWLQ 300
Db 241 RNOQSDPYKPIEIVNKTLPVATITLLSLVYLAGLLAAAYQLYGTYKRRFPPELFTWLQ 300
QY 301 CRKOLGLSFFFAVAVAYSLCLPMRSEERYFLNMAVYQOVHANIENSMNEEYVRLEMY 360
Db 301 CRKOLGLSFFFAVAVAYSLCLPMRSEERYFLNMAVYQOVHANIENSMNEEYVRLEMY 360
QY 361 ISFGIMSLGLSLAVTISPEVSNAIWMRERSFIQ 395
Db 361 ISFGIMSLGLSLAVTISPEVSNAIWMRERSFIQ 395

RESULT 12
ABG12306 standard; protein; 576 AA.

XX ABG12306;
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #12297.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS76493.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 20; SEQ ID NO 42665; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG030377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 576 AA;

Query Match 83.9%; Score 1972.5; DB 4; Length 576;
Best Local Similarity 87.4%; Pred. No. 6.7e-204;
Matches 390; Conservative 0; Mismatches 1; Indels 55; Gaps 1;

QY 1 MESISMWSPKSLSETCPNGINGIKDARKVTAVYIGSGDFAKSLTIRLRIGCHVYIGS 60
DB 1 MESISMWSPKSLSETCPNGINGIKDARKVTAVYIGSGDFAKSLTIRLRIGCHVYIGS 60
QY 61 RNKPFASFPFHVDVTHEDALTKNTIIFVAIHREHYTSLMDLRLHLLVKGKLLIDVSNM 120
DB 61 RNKPFASFPFHVDVTHEDALTKNTIIFVAIHREHYTSLMDLRLHLLVKGKLLIDVSNM 120
QY 121 RINQYPSNAEYIASLPDLSLVKGFNVVSAMALQGPDAARQYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYIASLPDLSLVKGFNVVSAMALQGPDAARQYICSNNIQARQVIE 180
QY 181 LARQNLFIPIIDGSLSSAREIENLPRLFTLRGPPVVAISLATFPFLYSFRDVIHPYA 240
DB 181 LARQNLFIPIIDGSLSSAREIENLPRLFTLRGPPVVAISLATFPFLYSFRDVIHPYA 240
QY 241 RNQOSDFYKIPRIEIVNKTLPVATITLSLVYAGLAAAYQLYGKTKRRFPMLWTWQ 300
DB 241 RNQOSDFYKIPRIEIVNKTLPVATITLSLVYAGLAAAYQLYGKTKRRFPMLWTWQ 300
QY 301 CRKQGLISFFPAMVHVAISLCIPKRSERYFLNMAVQCHANIENSMNEEVRIMY 360
DB 301 CRKQGLISFFPAMVHVAISLCIPKRSERYFLNMAVQCHANIENSMNEEVRIMY 360
QY 361 ISFGIMSLGLSLAATSIPIVSNALMNRBFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGLSLAATSIPIVSNALMNRBFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
QY 421 EBYRYFTPPNFTALVLPISIVLDL 446
DB 421 EBYRYFTPPNFTALVLPISIVLDL 446
QY 461 EBYRYFTPPNFTALVLPISIVLDL 491
DB 461 EBYRYFTPPNFTALVLPISIVLDL 491

RESULT 13

ABG00113 standard; protein; 1273 AA.

ABG00113; 13-FEB-2002 (first entry)

Novel human diagnostic protein #104.

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HISE-) HISEQ INC.

PI Drmanac RT, Liu C, Tang YT;
XX WPI, 2001-639362/73.
DR N-PSDB; AA64300.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

Claim 20; SEQ ID NO 30472; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnosis as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG030377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1273 AA;

Query Match 83.9%; Score 1972.5; DB 4; Length 1273;
Best Local Similarity 87.4%; Pred. No. 2.3e-203;
Matches 390; Conservative 0; Mismatches 1; Indels 55; Gaps 1;

QY 1 MESISMWSPKSLSETCPNGINGIKDARKVTAVYIGSGDFAKSLTIRLRIGCHVYIGS 60
DB 73 MESISMWSPKSLSETCPNGINGIKDARKVTAVYIGSGDFAKSLTIRLRIGCHVYIGS 132
QY 61 RNKPFASFPFHVDVTHEDALTKNTIIFVAIHREHYTSLMDLRLHLLVKGKLLIDVSNM 120
DB 133 RNKPFASFPFHVDVTHEDALTKNTIIFVAIHREHYTSLMDLRLHLLVKGKLLIDVSNM 192
QY 121 RINQYPSNAEYIASLPDLSLVKGFNVVSAMALQGPDAARQYICSNNIQARQVIE 180
DB 121 RINQYPSNAEYIASLPDLSLVKGFNVVSAMALQGPDAARQYICSNNIQARQVIE 180
QY 181 LARQNLFIPIIDGSLSSAREIENLPRLFTLRGPPVVAISLATFPFLYSFRDVIHPYA 240
DB 253 LARQNLFIPIIDGSLSSAREIENLPRLFTLRGPPVVAISLATFPFLYSFRDVIHPYA 312
QY 241 RNQOSDFYKIPRIEIVNKTLPVATITLSLVYAGLAAAYQLYGKTKRRFPMLWTWQ 300
DB 313 RNQOSDFYKIPRIEIVNKTLPVATITLSLVYAGLAAAYQLYGKTKRRFPMLWTWQ 372
QY 301 CRKQGLISFFPAMVHVAISLCIPKRSERYFLNMAVQCHANIENSMNEEVRIMY 360
DB 373 CRKQGLISFFPAMVHVAISLCIPKRSERYFLNMAVQCHANIENSMNEEVRIMY 411
QY 361 ISFGIMSLGLSLAATSIPIVSNALMNRBFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420
DB 412 -----QSTLGYVALLISTFHVLIYGMKRAFE 437
QY 421 EBYRYFTPPNFTALVLPISIVLDL 446
DB 438 EBYRYFTPPNFTALVLPISIVLDL 463

RESULT 14
AAE02841
ID AAE02841 standard; protein; 450 AA.
XX
AC AAE02841;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human STEAP-2 protein, alternative version.
XX
KW Human; cytostatic; antiproliferative; vaccine; gene therapy;
XX six transmembrane epithelial antigen of the prostate-2; STEAP-2;
KW chromosome 7q21; cancer; prostate; colon; bladder; lung; ovarian;
KW pancreatic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Misc-difference 335..336
FT /note= "Encoded by GACTGAGCT"
FT Misc-difference 375..376
FT /note= "Encoded by AGATGAGT"
FT Misc-difference 415..416
FT /note= "Encoded by GAGTAAGGC"
FT Misc-difference 445..446
FT /note= "Encoded by ACATGAGT"
FT Misc-difference 448..449
FT /note= "Encoded by AATTATTC"
XX
PN WO200140276-A2.
XX
PD 07-JUN-2001.
XX
PE 06-DEC-2000; 2000WO-US033040.
XX
PR 06-DEC-1999; 99US-00455486.
XX
PA (UROC-) UROGENESYS INC.
XX
XX Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC;
XX Faris M, Jakobovits A;
XX WPI; 2001-367804/38.
XX DR N-PSDB; AAD07072.
XX
PT New STEAP (six transmembrane epithelial antigen of the prostate)
PT proteins, expressed in human cancers, useful for detecting and treating
PT cancer.
XX
PS Claim 1; Page 175-176; 187pp; English.
XX
XX The present sequence is an alternative version of human six transmembrane
XX epithelial antigen of the prostate (STEAP)-2 protein. STEAP is a member
XX of cell surface serpentine transmembrane antigens. STEAP-2 gene is
XX located on chromosome 7q21 and is used in gene therapy. Inhibiting the
XX development or progression of a cancer (eg prostate, colon, bladder,
XX lung, ovarian and pancreatic) expressing STEAP or inhibiting growth or
XX killing cells expressing STEAP in a patient, comprises administering a
XX vaccine composition to the patient. Treating a patient with a cancer that
XX expresses STEAP, or inhibiting growth or killing cells expressing STEAP,
XX comprises administering to the patient a vector encoding single chain
XX monoclonal antibody that comprises the variable domains of the heavy and
XX light chains of the monoclonal antibody that specifically binds to STEAP.
XX such that the vector delivers the single chain monoclonal antibody coding
XX sequence to the cancer cells and the encoded single chain monoclonal
XX antibody is expressed intracellularly. Note: This sequence is stated to
XX be the same as that being shown as SEQ ID NO:8 (AAE02781) in figure 9A-9C
XX of the specification. However the present sequence lacks several amino
XX acids at its N-terminal end and has additional amino acids at its C-
XX terminal end when compared with the sequence shown in figure 9A-9C
XX
SQ Sequence 450 AA;

Query Match 73.9%; Score 1738; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. NO. 1,1e-178;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 119 NMRINQYPSNABYASLPDLSLYKGFNVVSAMALQGPXASRQVYICNNIQAQQV 178
DB 1 NMRINQYPSNABYASLPDLSLYKGFNVVSAMALQGPXASRQVYICNNIQAQQV 60
QY 179 IILARQLNPIPIDGLSLSSAREIENLPRLFTWMPGVVVAISLAFELSPVADVHP 238
DB 61 IILARQLNPIPIDGLSLSSAREIENLPRLFTWMPGVVVAISLAFELSPVADVHP 120
QY 239 YARNQOSDFYKPIPIEIVKTLPIVAITLISLVYLAGLAAAYQLYGTRKRFPMLETW 298
DB 121 YARNQOSDFYKPIPIEIVKTLPIVAITLISLVYLAGLAAAYQLYGTRKRFPMLETW 180
QY 299 LOCRKQLGLSFFPMVAVYSLCLPMRSEERYFLNMAVQOVANINSENNEEVRLE 358
DB 181 LOCRKQLGLSFFPMVAVYSLCLPMRSEERYFLNMAVQOVANINSENNEEVRLE 240
QY 359 MYISFGIMSLGLSLAATVTSIPSVSNALNMRFSFIOSTLGVALLISTFHYLYGWTRA 418
DB 241 MYISFGIMSLGLSLAATVTSIPSVSNALNMRFSFIOSTLGVALLISTFHYLYGWTRA 300
QY 419 FEEYRYRFTYPPNFVLAIVLPISVILDLQLCRYPD 454
DB 301 FEEYRYRFTYPPNFVLAIVLPISVILDLQLCRYPD 336
RESULT 15
ABP62883
ID ABP62883 standard; protein; 1082 AA.
XX
AC ABP62883;
XX
DT 14-OCT-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 320.
XX
XX Human; vulnary; dermatological; neuroprotective; nootropic; cancer;
XX antiparkinsonian; immunostimulant; cyostatic; immunosuppressive;
XX antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;
XX burn; central nervous system disorder; Alzheimer's disease;
XX Parkinson's disease; Huntington's disease; immune disorder;
XX autoimmune disorder; multiple sclerosis; diabetes; allergy.
XX
OS Homo sapiens.
XX
XX WO200218424-A2.
XX
PD 07-MAR-2002.
XX
XX 31-AUG-2001; 2001WO-US027093.
XX
PR 01-SEP-2000; 2000US-00654935.
XX
XX (HYSE-) HYSQ INC.
XX
XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;
XX Zhao QA, Wang D, Liu C, Dirmann RT, Wehrman T;
XX WPI; 2002-583321/62.
XX DR N-PSDB; ABQ93362.
XX
XX New polynucleotide and polypeptides, useful for treatment and diagnosis
XX of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral
XX sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple
XX sclerosis, diabetes and allergies.
XX
PS Claim 20; SEQ ID NO 320; 284pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (i) comprising one of
XX 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising

CC administering to a mammalian subject a composition comprising the protein
CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).
CC (II), (III) and (III) are useful for diagnostic evaluation of disorders.
CC (I) is useful for gene therapy of diseases and (II) can be used for
CC therapeutic treatment. Diseases that may be treated include wound healing
CC and tissue repair, burns, central nervous system disorders (e.g.
CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral
CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple
CC sclerosis, diabetes and allergies. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 1082 AA;